(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
11 January 2001 (11.01.2001)

PCT

(10) International Publication Number
WO 01/02828 A2

(51) International Patent Classification[7]: G01N

(21) International Application Number: PCT/US00/18657

(22) International Filing Date: 7 July 2000 (07.07.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/142,720	7 July 1999 (07.07.1999)	US
Not furnished	7 July 2000 (07.07.2000)	US

(71) Applicant *(for all designated States except US)*: **TULARIK INC.** [US/US]; Two Corporate Drive, South San Francisco, CA 94080 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **ERSTER, Susan** [US/US]; 2 Melville Court, Stony Brook, NY 11790 (US). **POWERS, Scott** [US/US]; 6 Westbrook Court, Greenlawn, NY 11740 (US).

(74) Agents: **GARRETT-WACKOWSKI, Eugenia** et al.; Townsend Townsend And Crew LLP, Two Embarcadero Center, Eighth Floor, San Francisco, CA 94111 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

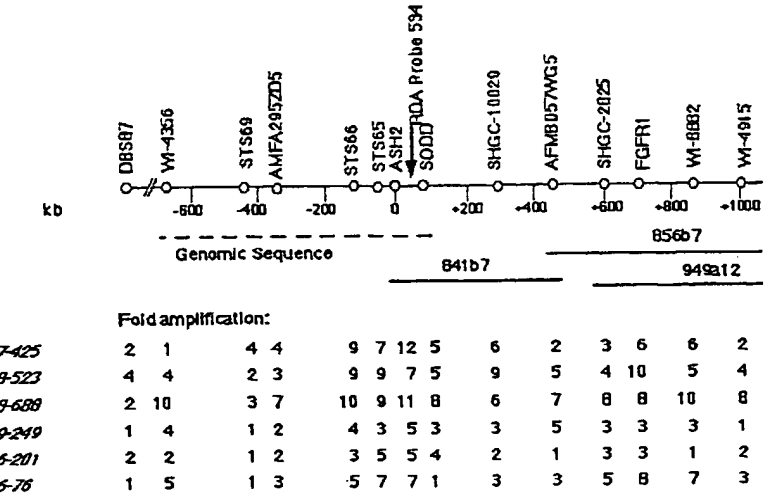
(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: DIAGNOSIS OF CANCER BY DETECTING ASH2 POLYPEPTIDES OR POLYNUCLEOTIDES



(57) Abstract: The present invention provides methods for assessing the presence of a cancer in an animal. The methods provided herein are based on the surprising discovery that the ASH2 gene is amplified and/or overexpressed in multiple types of cancer cells. Accordingly, it has been discovered that the detection of overexpression or amplification of ASH2 in a biological sample from an animal allows the diagnosis of cancer in the sample. In addition, the present invention provides numerous methods for treating cancer, including by determining the most effective course of anti-cancer therapy based on ASH2 levels as well as by inhibiting the proliferation of cancer cells by inhibiting or reducing ASH2 polypeptide or polynucleotide levels.

WO 01/02828 A2

DIAGNOSIS OF CANCER BY DETECTING ASH2 POLYPEPTIDES OR POLYNUCLEOTIDES

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH

Not Applicable

STATEMENT AS TO RIGHTS TO INVENTIONS MADE UNDER FEDERALLY SPONSORED RESEARCH AND DEVELOPMENT

Not Applicable

BACKGROUND OF THE INVENTION

Despite years of research into its causes and potential treatments, cancer remains the second leading cause of death in the United States. Currently, more than 500 thousand Americans die of cancer each year, and more than 1.2 million new cancer cases will be diagnosed in 1999. Although some progress has been made towards understanding the causes of cancer, a major need remains for new tools for the diagnosis and treatment of cancer.

Of the many types of cancer, epithelial cancers are among the most prevalent and deadly. For example, about 175,000 new invasive cases of breast cancer are expected to occur among women in the United States during 1999, and more than 43,000 women will die of this disease. In addition, an estimated 129,400 cases of colorectal cancer are expected to occur in 1999, including 94,700 cases of colon cancer and 34,700 cases of rectal cancer. Colorectal cancer is the third most common cancer in men and women. An estimated 56,600 deaths (47,900 from colon cancer, 8,700 from rectal cancer) are expected to occur in 1999.

Cancer is a genetic disease of single cell origin caused by the accumulation of inherited and acquired mutations in specific cancer genes which have normal cellular functions, but which contribute to cancer when mutated. Mutations that unlock the cancer-causing potential of cancer genes include, for example, gene amplification, *i.e.*, where a specific chromosomal region (including the cancer gene) undergoes a relative increase in DNA copy number, so that more cancer gene copies are present and a correspondingly higher amount of mRNA and protein is produced, causing deleterious effects. The

identification of such amplified (or overexpressed) genes may provide important tools for the diagnosis and/or treatment of cancer. To date, several such overexpressed genes have been identified, such as *HER2* and *MYC*. These genes, however, are not involved in all types of cancers, and many cancers contain amplified regions where the underlying cancer gene is not known.. Clearly, a large number of genes that are amplified and/or overexpressed in cancer cells remain to be identified.

One particularly effective method of identifying amplified regions of the genome is through representational difference analysis, or RDA (Lisitsyn *et al.*, (1993) *Science* 259:946-951; U.S. Patent Nos. 5,436,142 and 5,501,964). Briefly, RDA involves several steps, including isolating a "representative" population of nucleic acid from a cell or cell type or interest, *e.g.*, a cancer cell. This "representative" population contains significantly less complexity than that found in the overall genome. Using a simultaneous hybridization and amplification step to enrich for those sequences not present in a control, or "tester" cell or cell population, RDA allows for the identification of sequences that are over-represented in the cell type of interest.

In the fruitfly *Drosophila melanogaster*, the ASH2 gene is a member of the "trithorax" group of genes whose products function to maintain active transcription of homeotic selector genes. Mutations in Drosophila ASH2 result in homeotic transformations as well as a variety of pattern formation defects (*see, e.g.*, *Genetics* (1996) 144(2):621-633; see also the entry for ASH2 at Flybase, a Drosophila gene database, at http://flybase.bio.indiana.edu/). Homologs of ASH2 have been identified in organisms other than Drosophila, such as humans, mice, and fish, but no function has been attributed to any of these non-Drosophila ASH2 homologs.

The present invention is based on the surprising discovery that ASH2 polynucleotide sequences are amplified and/or overexpressed in many types of cancer cells in mammals. As described *infra*, this invention thus provides novel and badly needed diagnostic, prognostic, and therapeutic tools for many types of cancers.

SUMMARY OF THE INVENTION

As such, the present invention provides methods for diagnosing and treating cancer in an animal, *e.g.*, a human. The methods typically involve detecting the level of ASH2 nucleic acid or protein in a biological sample taken from the animal. The presence of ASH2 nucleic acid and/or protein in the sample at a level that is greater than that expected

for a non-cancerous sample is indicative of cancer in the sample, or in the animal from which the sample is derived.

In numerous embodiments, the presence of a cancer in an animal will be assessed by detecting the presence or absence of a diagnostic presence of ASH2 polypeptide or polynucleotide in a biological sample taken from the animal. In such embodiments, the assay used to detect ASH2 in the sample will often be performed under conditions that would not detect ASH2 polypeptide or polynucleotide in a sample that is not cancerous. In such embodiments, a detection of a diagnostic presence of ASH2 indicates the presence of cancer in the animal. In preferred embodiments, a diagnostic presence represents at least about a 2, 5, 10, or greater fold increase in the ASH2 polypeptide or polynucleotide in the biological sample compared to a level expected in a sample from a control, cancer-free animal.

In numerous embodiments, the present invention provides methods for monitoring the efficacy of a cancer treatment. In such methods, a level of ASH2 polypeptide or polynucleotide in a biological sample from an animal undergoing treatment for cancer is detected. In such embodiments, a reduced level of ASH2 polypeptide or polynucleotide in the biological sample compared to a level in a sample taken from the animal prior to, or earlier in, the treatment, indicates that the treatment is efficacious.

In numerous embodiments, the methods provided herein can be used to treat cancer. In typical embodiments, the presence or absence of a diagnostic presence of ASH2 polynucleotide or polypeptide is detected in a biological sample taken from an animal with cancer. In such embodiments, the detection of a diagnostic presence of ASH2, indicating the presence of cancer in the biological sample, is followed by the administration of one or more types of cancer therapy. In such embodiments, the detection of a presence or absence of a diagnostic presence of ASH2 polynucleotide or polypeptide is often repeated after, or during, the administration of the cancer therapy, thereby allowing a determination of the efficacy of the cancer therapy.

Also provided herein are methods of decreasing the proliferation of a cell with an elevated level of ASH2 activity. In such embodiments, the proliferation is decreased using an inhibitor of ASH2 activity, such as antisense polynucleotides, ribozymes, antibodies, dominant negative ASH2 inhibitors, and small molecule inhibitors of ASH2 activity.

In numerous embodiments, the present methods are performed by detecting an ASH2 gene, an ASH2 mRNA, an ASH2 polypeptide or ASH2 protein activity.

In preferred embodiments of this invention, the animal is a mammal, *e.g.*, a primate, canine, feline, murine, bovine, equine, ovine, porcine, lagomorph, *etc.* In particularly preferred embodiments, the animal is a human.

In preferred embodiments of the invention, a biological sample used for detection of ASH2 is a sample selected from the group consisting of tissue biopsy, blood sample, buccal scrape, saliva, nipple discharge, urine, *etc.*

The present methods can be used to diagnose, determine the prognosis for, or treat, any of a number of types of cancers. In preferred embodiments, the cancer is an epithelial cancer, *e.g.*, breast, lung, colorectal, prostate, kidney, stomach, bladder, or ovarian cancer, or any cancer of the gastrointestinal tract.

The present invention also provides ASH2 polynucleotides and polypeptides. The ASH2 polynucleotides and polypeptides provided herein represent novel splice variants that lack N-terminal and exon sequences found in previously described ASH2 sequences. In addition, the ASH2 polypeptide sequences provided herein comprise novel subsequences not described previously. For example, the present invention provides ASH2 polynucleotides, and polypeptides encoded thereby, that encode ASH2 polypeptides including novel subsequences that are at least about 70% identical to the sequences shown as SEQ ID NO:3 or SEQ ID NO:9, or which lack the sequence shown as SEQ ID NO:4.

Kits for practicing the present invention are also provided.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a physical map of the 8p11-p12 amplified chromosomal region, indicated by the solid line marked with solid circles.

DETAILED DESCRIPTION OF THE INVENTION AND PREFERRED EMBODIMENTS

I. Introduction

The present invention provides methods for assessing the presence of cancer in a biological sample taken from an animal. The methods provided herein are based on the surprising discovery that the ASH2 gene, forms of which have been previously identified in mammals, but with no known function, is amplified and/or overexpressed in numerous types of cancers. Accordingly, it has been discovered that the detection of overexpression or amplification of ASH2 polynucleotides or polypeptides in a biological sample from an animal is diagnostic of cancer in the animal. Detection of ASH2 in a sample also provides a

means for monitoring the efficacy of a cancer treatment. In addition, the present invention provides numerous methods for treating cancer, including determining the most effective course of anti-cancer therapy based on ASH2 levels, and methods of inhibiting the proliferation of cancer cells by reducing ASH2 polypeptide or polynucleotide levels or inhibiting ASH2 protein activity. Also provided are novel ASH2 polynucleotide and polypeptide sequences.

II. Definitions

The phrase "detecting a cancer" refers to the ascertainment of the presence or absence of cancer in an animal. "Detecting a cancer" can also refer to obtaining indirect evidence regarding the likelihood of the presence of cancerous cells in the animal. Detecting a cancer can be accomplished using the methods of this invention alone or in combination with other methods, and can be aided by other information regarding the state of health of the animal.

A "cancer" in an animal refers to the presence of cells possessing characteristics typical of cancer-causing cells, such as uncontrolled proliferation, immortality, metastatic potential, rapid growth and proliferation rate, and certain characteristic morphological features. Often, cancer cells will be in the form of a tumor, but such cells may exist alone within an animal, or may circulate in the blood stream as independent cells, such as leukemic cells.

An "animal" refers to a member of the kingdom Animalia, characterized by multicellularity, the possession of a nervous system, voluntary movement, internal digestion, *etc.* An "animal" can be a human or any other mammal, including non-human primates, canines, felines, murines, bovines, ovines, equines, porcines, and lagomorphs.

"Providing a biological sample" means to obtain a biological sample for use in the methods described in this invention. Most often, this will be done by removing a sample of cells from an animal, but can also be accomplished by using previously isolated cells (*e.g.*, isolated by another person, at another time, and/or for another purpose), or by performing the methods of the invention *in vivo*.

A "control" sample refers to a sample of biological material representative of healthy, cancer-free animals. The level of ASH2 in a control sample is desirably typical of the general population of normal, cancer-free animals. This sample can be removed from an animal expressly for use in the methods described in this invention, or can be any biological material representative of normal, cancer-free animals. A control sample can also refer to an

established level of ASH2, representative of the cancer-free population, that has been previously established based on measurements from normal, cancer-free animals. If a detection method is used that only detects ASH2 when a level higher than that typical of a normal, cancer-free animal is present, *i.e.*, an immunohistochemical assay giving a simple positive or negative result, this is considered to be assessing the ASH2 level in comparison to the "control" level, as the "control" level is inherent in the assay.

A level of ASH2 polypeptide or polynucleotide that is "expected" in a control sample refers to a level that represents a typical, cancer-free sample, and from which an elevated, or diagnostic, presence of ASH2 polypeptide or polynucleotide can be distinguished. Preferably, an "expected" level will be controlled for such factors as the age, sex, medical history, *etc.* of the animal, as well as for the type of biological sample being tested.

An "increased," or "elevated," level of ASH2 refers to a level of ASH2 polynucleotide or polypeptide, that, in comparison with a control level of ASH2, is detectably higher. The method of comparison can be statistical, using quantified values for the level of ASH2, or can be compared using non-statistical means, such as by a visual, subjective assessment by a human.

A "diagnostic presence" of ASH2 polynucleotides or polypeptides refers to any amount of ASH 2 in a biological sample, detected using any method, that represents an increase over a control level.

An "ASH2 polynucleotide" or "ASH2 nucleic acid" is a DNA or RNA sequence of at least about 50 nucleotides that is at least about 70% identical, preferably at least about 80% or more, identical over a region of at least about 50, 100, 200, 500, or more nucleotides to one or more ASH2 polynucleotide sequences (*see, e.g.*, GenBank Accession Nos: AF056717, AB020982, AF056718, AB020983, SEQ ID NO:1, SEQ ID NO:5, SEQ ID NO:7, and others). An "ASH2 polynucleotide" can comprise naturally occuring nucleotides, or any derivative or analog thereof, *e.g.*, labeled or modified deoxyribo- or ribonucleotides. The term "ASH2 polynucleotide" can refer to a mutated copy of any of the above sequences, or a fragment thereof. An "ASH2 polynucleotide" can refer to a natural sequence derived from an ASH2 polynucleotide from any organism, or can be a sequence designed *de novo*.

An "ASH2 protein" or "ASH2 polypeptide" refers to a polypeptide of at least about 20 amino acids that is typically about 70% identical, preferably at least about 80%, more preferably at least about 90%, or more, identical over a region of at least about 20, 50, 100 or more amino acids, to one or more of the ASH2 polypeptide sequences (*see, e.g.*,

GenBank Accession Nos: AF056717, AB020982, AF056718, AB020983, or SEQ ID NO:2, SEQ ID NO:6, SEQ ID NO:8), or to any derivative, variant, mutant, or fragment thereof. An "ASH2 protein" can comprise naturally occurring or synthetic amino acids, *e.g.*, labeled or otherwise modified amino acids or amino acid analogs. An "ASH2 protein" will typically contain one or more characteristic protein motifs, *e.g.*, zinc finger motif, any of which can be used independently of other elements normally present in a full-length ASH2 protein. An "ASH2 protein" can refer to a natural sequence derived from an ASH2 polypeptide from any organism, or to a sequence designed *de novo*.

The "level of ASH2 mRNA" in a biological sample refers to the amount of mRNA transcribed from an ASH2 gene, or encoding an ASH2 polypeptide, that is present in a cell or a biological sample. The mRNA generally encodes a functional ASH2 protein, although mutations or microdeletions may be present that alter or eliminate the function of the encoded protein. A "level of ASH2 mRNA" need not be quantified, but can simply be detected, *e.g.*, a subjective, visual detection by a human, with or without comparison to a level from a control sample or a level expected of a control sample.

The "level "of ASH2 "protein," or "polypeptide" in a biological sample refers to the amount of polypeptide translated from an ASH2 mRNA, or to the amount of ASH2 protein as defined *supra*, that is present in a cell or biological sample. The polypeptide may or may not have ASH2 protein activity, *i.e.*, it may or may not bind to one or more DNA sequences or heterologous proteins. A "level of ASH2 protein" need not be quantified, but can simply be detected, *e.g.*, a subjective, visual detection by a human, with or without comparison to a level from a control sample or a level expected of a control sample.

A "full length" ASH2 protein or nucleic acid refers to an ASH2 polypeptide or polynucleotide sequence, or a variant thereof, that contains all of the elements normally contained in one or more naturally occurring, wild type ASH2 polynucleotide or polypeptide sequences.

When a quantified level of ASH2 falls outside of a given confidence interval for a normal level of ASH2, the difference between the two levels is said to be "statistically significant." If a test value falls outside of a given confidence interval for a normal level of ASH2, it is possible to calculate the probability that the test value is truly abnormal and does not just represent a normal deviation from the average. In the present invention, a difference between a test sample and a control can be termed "statistically significant" when the probability of the test sample being abnormal can be any of a number of values, including

0.15, 0.1, 0.05, and 0.01. Numerous sources teach how to assess statistical significance, such as Freund, J.E. (1988) *Modern elementary statistics*, Prentice-Hall.

When a level of ASH2 in a sample is said to be "at least about 2 fold greater" than the level of ASH2 in another sample or in a control, this signifies that the level of ASH2 in the first sample is at least about 2 times the value of ASH2 in the second sample or to a control value.

When a level of ASH2 in a sample is said to be "at least about 5 fold greater" than the level of ASH2 in another sample or in a control, this signifies that the level of ASH2 in the first sample is at least about 5 times the value of ASH2 in the second sample or to a control value.

The terms "identical" or percent "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same, when compared and aligned for maximum correspondence, as measured using one of the following sequence comparison algorithms or by visual inspection.

The phrase "substantially identical," in the context of two nucleic acids or polypeptides, refers to two or more sequences or subsequences that have at least 60%, preferably 80%, most preferably 90-95% nucleotide or amino acid residue identity, when compared and aligned for maximum correspondence, as measured using one of the following sequence comparison algorithms or by visual inspection. Preferably, the substantial identity exists over a region of the sequences that is at least about 50 residues in length, more preferably over a region of at least about 100 residues, and most preferably the sequences are substantially identical over at least about 150 residues. In a most preferred embodiment, the sequences are substantially identical over the entire length of the coding regions.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are input into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. The sequence comparison algorithm then calculates the percent sequence identity for the test sequence(s) relative to the reference sequence, based on the designated program parameters.

Optimal alignment of sequences for comparison can be conducted, *e.g.*, by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444

(1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by visual inspection (*see generally, Current Protocols in Molecular Biology*, F.M. Ausubel *et al.*, eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (1995 Supplement) (Ausubel)).

Examples of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al.* (1990) *J. Mol. Biol.* 215: 403-410 and Altschuel *et al.* (1977) *Nucleic Acids Res.* 25: 3389-3402, respectively. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al, supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (*see* Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)).

In addition to calculating percent sequence identity, the BLAST algorithm also performs a statistical analysis of the similarity between two sequences (*see, e.g.*, Karlin & Altschul, *Proc. Nat'l. Acad. Sci. USA* 90:5873-5787 (1993)). One measure of similarity

provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

A further indication that two nucleic acid sequences or polypeptides are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the polypeptide encoded by the second nucleic acid, as described below. Thus, an ASH2 polypeptide is typically substantially identical to a second ASH2 polypeptide, for example, where the two peptides differ only by conservative substitutions. Another indication that two ASH2 nucleic acid sequences are substantially identical is that the two molecules hybridize to each other under stringent conditions, as described below.

The phrase "hybridizing specifically to" refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence under stringent conditions when that sequence is present in a complex mixture (*e.g.*, total cellular) DNA or RNA. The term "stringent conditions" refers to conditions under which a probe will hybridize to its target subsequence, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. Generally, stringent conditions are selected to be about 5°C lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength, pH, and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. (As the target sequences are generally present in excess, at Tm, 50% of the probes are occupied at equilibrium). Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (*e.g.*, 10 to 50 nucleotides) and at least about 60°C for long probes (*e.g.*, greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide.

III. Detecting ASH2 Polynucleotides and Polypeptides

The present invention is based on the discovery that a diagnosis or prognosis of cancer can be determined by detecting elevated levels of ASH2 polypeptides or polynucleotides in a biological sample from an animal. As described below, any of a number of methods to detect the presence and/or levels of ASH2 can be used. An ASH2 polynucleotide level can be detected by detecting the presence of any ASH2 DNA or RNA, including ASH2 genomic DNA, mRNA, and cDNA. An ASH2 polypeptide can be detected by detecting an ASH2 polypeptide itself, or by detecting ASH2 protein activity, *e.g.*, DNA or protein binding activity, transcriptional regulation, *etc.* Detection can involve quantification of the level of ASH2 (*e.g.*, gDNA, cDNA, mRNA, protein, or protein activity), or, alternatively, can be a qualitative assessment of the level, or of the presence or absence, of ASH2, in particular in comparison with a control level. Any of a number of methods to detect any of the above can be used, as described *infra*. Such methods include, for example, hybridization, amplification, and other assays.

In certain embodiments, a level of ASH2 in a biological sample will be compared with a control sample taken from a cancer-free animal, or, preferably, with a value expected for a sample taken from a cancer-free animal. In a particularly preferred embodiment, an assay will be performed under conditions where only a higher than normal amount of ASH2 polynucleotide or polypeptide will be detectable in the assay. As a result, an elevated level of ASH2 can be detected in a sample using a simple assay giving a simple, positive or negative result, with no need for quantification of ASH2 levels or a direct comparison with a control sample.

In certain embodiments, the level of ASH2 polynucleotide, polypeptide, or protein activity will be quantified. In such embodiments, the difference between an elevated level of ASH2 and a normal, control level will preferably be statistically significant. In preferred embodiments, an elevated level of ASH2 polynucleotide, polypeptide, and/or protein activity will be at least about 2, 5, 10, or more fold greater than a control level.

Typically, the ASH2 polynucleotides or polypeptides detected herein will be at least about 70% identical, and preferably 80% or more identical, over a region of at least about 50, 100, 200, or more nucleotides, or 20, 50, 100, or more amino acids, to SEQ ID NO:1, 5, or 7, or SEQ ID NO:2, 6, or 8, or to one or more sequences available, *e.g.*, from GenBank (*see, e.g.*, GenBank Accession Nos: AF056717, AB020982, AF056718, AB020983, U73809, AB022785, and others). Such polynucleotides or polypeptides can

represent any of the ASH2 variants described herein, *e.g.*, splice variants, and can indicate functional or non-functional forms of ASH2, or any variant, derivative, or fragment thereof.

A. Providing a Biological Sample

Typically, the level and/or presence of ASH2 polynucleotides or polypeptides will be detected in a biological sample. A "biological sample" refers to a cell or population of cells or a quantity of tissue or fluid from an animal. Most often, the sample has been removed from an animal, but the term "biological sample" can also refer to cells or tissue analyzed *in vivo*, *i.e.*, without removal from the animal. Typically, a "biological sample" will contain cells from the animal, but the term can also refer to non-cellular biological material, such as non-cellular fractions of blood, saliva, or urine. Numerous types of biological samples can be used in the present invention, including, but not limited to, a tissue biopsy, blood sample, a buccal scrape, a saliva sample, or a nipple discharge.

As used herein, a "tissue biopsy" refers to an amount of tissue removed from an animal for diagnostic analysis. In a patient with cancer, tissue may be removed from a tumor, allowing the analysis of cells within the tumor. "Tissue biopsy" can refer to any type of biopsy, such as needle biopsy, fine needle biopsy, surgical biopsy, *etc.* A "buccal scrape" refers to a sample of cells removed from the inner lining of the mouth. A "nipple discharge" refers to fluid originating from a nipple, which may contain cancerous cells or may contain elevated levels of ASH2 polypeptide indicating the presence of cancerous cells in the breast.

B. Detection of Copy Number

In one embodiment, the presence of cancer is evaluated by determining the copy number of ASH2 genes. The "copy number of ASH2 genes" refers to the number of DNA sequences in a cell encoding an ASH2 protein. Generally, for a given autosomal gene, an animal has two copies of each gene. The copy number can be increased, however, by gene amplification or duplication, *e.g.*, in cancer cells, or reduced by deletion. Methods of evaluating the copy number of a particular gene are well known to those of skill in the art, and include, *inter alia*, hybridization and amplification based assays.

1. Hybridization-based Assays

Any of a number of hybridization based assays can be used to detect the copy number of ASH2 genes in the cells of a biological sample. One such method is by Southern

Blot. In a Southern Blot, genomic DNA is typically, fragmented, separated electrophoretically, and hybridized to an ASH2 specific probe. Comparison of the intensity of the hybridization signal from the probe for the target region with control probe signal from analysis of normal genomic DNA (*e.g.*, a non-amplified portion of the same or related cell, tissue, organ, *etc.*) provides an estimate of the relative ASH2 copy number.

An alternative means for determining the copy number of ASH2 genes in a sample is *in situ* hybridization, *e.g.*, fluorescence *in situ* hybridization, or FISH. *In situ* hybridization assays are well known (*e.g.*, Angerer (1987) *Meth. Enzymol* 152: 649). Generally, *in situ* hybridization comprises the following major steps: (1) fixation of tissue or biological structure to be analyzed; (2) prehybridization treatment of the biological structure to increase accessibility of target DNA, and to reduce nonspecific binding; (3) hybridization of the mixture of nucleic acids to the nucleic acid in the biological structure or tissue; (4) post-hybridization washes to remove nucleic acid fragments not bound in the hybridization and (5) detection of the hybridized nucleic acid fragments.

The probes used in such applications are typically labeled, *e.g.*, with radioisotopes or fluorescent reporters. Preferred probes are sufficiently long, *e.g.*, from about 50, 100, or 200 nucleotides to about 1000 or more nucleotides, so as to specifically hybridize with the target nucleic acid(s) under stringent conditions.

In preferred embodiments, "comparative probe" methods, such as comparative genomic hybridization (CGH), are used to detect ASH2 gene amplification. In comparative genomic hybridization methods, a "test" collection of nucleic acids is labeled with a first label, while a second collection (*e.g.*, from a healthy cell or tissue) is labeled with a second label. The ratio of hybridization of the nucleic acids is determined by the ratio of the first and second labels binding to each fiber in the array. Differences in the ratio of the signals from the two labels, *e.g.*, due to gene amplification in the test collection, is detected and the ratio provides a measure of the ASH2 gene copy number.

Hybridization protocols suitable for use with the methods of the invention are described, *e.g.*, in Albertson (1984) *EMBO J.* 3: 1227-1234; Pinkel (1988) *Proc. Natl. Acad. Sci. USA* 85: 9138-9142; EPO Pub. No. 430,402; *Methods in Molecular Biology, Vol. 33: In Situ Hybridization Protocols,* Choo, ed., Humana Press, Totowa, NJ (1994), and elsewhere.

2. Amplification-based Assays.

In still another embodiment, amplification-based assays are used to measure an ASH2 copy number. In such assays, the ASH2 nucleic acid sequences act as a template

in an amplification reaction (*e.g.*, Polymerase Chain Reaction, or PCR). In a quantitative amplification, the amount of amplification product will be proportional to the amount of template in the original sample. Comparison to appropriate controls provides a measure of the copy number of the ASH2 gene. Methods of quantitative amplification are well known to those of skill in the art. Detailed protocols for quantitative PCR are provided, *e.g.*, in Innis *et al.* (1990) *PCR Protocols, A Guide to Methods and Applications*, Academic Press, Inc. N.Y.). The known nucleic acid sequence for ASH2 (*see, e.g.*, SEQ ID NO:1) is sufficient to enable one of skill to routinely select primers to amplify any portion of the gene.

In preferred embodiments, a TaqMan based assay is used to quantify ASH2 polynucleotides. TaqMan based assays use a fluorogenic oligonucleotide probe that contains a 5' fluorescent dye and a 3' quenching agent. The probe hybridizes to a PCR product, but cannot itself be extended due to a blocking agent at the 3' end. When the PCR product is amplified in subsequent cycles, the 5' nuclease activity of the polymerase, *e.g.*, AmpliTaq, results in the cleavage of the TaqMan probe. This cleavage separates the 5' fluorescent dye and the 3' quenching agent, thereby resulting in an increase in fluorescence as a function of amplification (*see*, for example, literature provided by Perkin-Elmer, *e.g.*, www2.perkin-elmer.com).

Other suitable amplification methods include, but are not limited to, ligase chain reaction (LCR) (*see,* Wu and Wallace (1989) *Genomics* 4: 560, Landegren *et al.* (1988) *Science* 241: 1077, and Barringer *et al.* (1990) *Gene* 89: 117), transcription amplification (Kwoh *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86: 1173), self-sustained sequence replication (Guatelli *et al.* (1990) *Proc. Nat. Acad. Sci. USA* 87: 1874), dot PCR, and linker adapter PCR, *etc.*

C. Detection of Gene Expression

In preferred embodiments, ASH2 levels are characterized by detecting ASH2 gene expression by virtue of levels of ASH2 mRNA or cDNA in a biological sample.

1. Detection of Gene Transcript

a) Direct hybridization-based assays

Methods of detecting and/or quantifying the level of ASH2 gene transcript (ASH2 mRNA or cDNA made therefrom) using nucleic acid hybridization techniques are

known to those of skill in the art (*see*, Sambrook *et al.*, (1989) *Molecular Cloning: A Laboratory Manual*, 2d Ed., vols 1-3, Cold Spring Harbor Press, New York).

For example, one method for evaluating the presence, absence, or quantity of ASH2 cDNA involves a Southern Blot as described above. Briefly, ASH2 mRNA is isolated using standard methods and reverse transcribed to produce cDNA. The cDNA is then optionally digested, run on a gel, and transferred to a membrane. Hybridization is then carried out using nucleic acid probes specific for ASH2 cDNA and detected using standard techniques (*see, e.g.*, Sambrook *et al.*, *supra*).

Similarly, a Northern transfer may be used to detect an mRNA directly. In a typical embodiment, mRNA is isolated from a given biological sample, electrophoresed to separate the mRNA species, and transferred from the gel to a nitrocellulose membrane. As with the Southern Blots, labeled ASH2 probes are used to identify and/or quantify the mRNA.

b) Amplification-based assays.

In another preferred embodiment, an ASH2 transcript (*e.g.*, ASH2 mRNA) can be measured using amplification-based methods (*e.g.*, PCR). In a preferred embodiment, a transcript level is assessed by using reverse transcription PCR (RT-PCR). RT-PCR methods are well known to those of skill (*see, e.g.*, Ausubel *et al.*, *supra*).

D. Detection of ASH2 Protein

ASH2 levels can also be detected and/or quantified by detecting or quantifying ASH2 polypeptide. ASH2 polypeptide can be detected and quantified by any of a number of means well known to those of skill in the art. These include analytic biochemical methods such as electrophoresis, capillary electrophoresis, high performance liquid chromatography (HPLC), thin layer chromatography (TLC), hyperdiffusion chromatography, and the like, or various immunological methods such as fluid or gel precipitin reactions, immunodiffusion (single or double), immunoelectrophoresis, radioimmunoassay (RIA), enzyme-linked immunosorbent assays (ELISAs), immunofluorescent assays, western blotting, and the like.

In a preferred embodiment, an ASH2 polypeptide is detected using an immunoassay such as an ELISA assay (*see, e.g.*, Crowther, John R. *ELISA Theory and Practice.* Humana Press: New Jersey, 1995). As used herein, an "immunoassay" is an assay that utilizes an antibody to specifically bind to the analyte (*i.e.*, the ASH2 polypeptide). The

immunoassay is thus characterized by detection of specific binding of an ASH2 polypeptide to an anti-ASH2 antibody.

In an immunoassay, ASH2 polypeptide can be detected and/or quantified using any of a number of well recognized immunological binding assays (*see, e.g.*, U.S. Patent Nos. 4,366,241; 4,376,110; 4,517,288; and 4,837,168). For a review of the general immunoassays, see also Asai (1993) *Methods in Cell Biology* Volume *37: Antibodies in Cell Biology,* Academic Press, Inc. New York; Stites & Terr (1991) *Basic and Clinical Immunology 7th Edition.*

Immunoassays typically rely on direct or indirect labeling methods to detect antibody-analyte binding. For example, an anti-ASH2 antibody can be directly labeled, thereby allowing detection. Alternatively, the anti-ASH2 antibody may itself be unlabeled, but may, in turn, be bound by a labeled third antibody specific to antibodies of the species from which the second antibody is derived. The second or third antibodies can also be modified with a detectable moiety, *e.g.*, as biotin, to which a third labeled molecule can specifically bind, such as enzyme-labeled streptavidin. Also, other antibody-binding molecules can be used, *e.g.*, labeled protein A or G (*see, generally* Kronval, *et al.* (1973) *J. Immunol.*, 111: 1401-1406, and Akerstrom (1985) *J. Immunol.*, 135: 2589-2542).

Immunoassays for detecting an ASH2 polypeptide can be competitive or noncompetitive. Noncompetitive immunoassays are assays in which the amount of captured analyte is directly measured. In a preferred embodiment, "sandwich" assays will be used, for example, wherein anti-ASH2 antibodies are bound directly to a solid substrate where they are immobilized. These immobilized antibodies then capture the ASH2 protein present in a test sample. The ASH2 thus immobilized is then bound by a labeling agent, such as a second anti-ASH2 antibody bearing a label.

In competitive assays, the amount of ASH2 protein present in a sample is measured indirectly, *e.g.*, by measuring the amount of added (exogenous) ASH2 displaced (or competed away) from an anti ASH2 antibody by ASH2 protein present in a sample. For example, a known amount of labeled ASH2 polypeptide is added to a sample and the sample is then contacted with an anti-ASH2 antibody. The amount of labeled ASH2 polypeptide bound to the antibody is inversely proportional to the concentration of ASH2 polypeptide present in the sample.

Any of a number of labels can be used in any of the immunoassays of this invention, including fluorescent labels, radioisotope labels, or enzyme-based labels, wherein

a detectable product of enzyme activity is detected (*e.g.*, peroxidase, alkaline phosphatase, β-galactosidase, *etc.*).

Antibodies for use in the various immunoassays described herein can be produced according to standard methods (*see, e.g.*, Harlow and Lane (1988) *Antibodies: A Laboratory Manual*, Cold Spring Harbor Press, NY.)

Such methods can also be used to distinguish between various amounts of ASH2 protein, as well as between various forms of ASH2 proteins. For example, ASH2 can exist in one or more splice variants, described *infra*, which can be distinguished in a variety of ways, *e.g.*, electrophoretic separation based on size, detection using variant specific antibodies, *etc.*

E. Detection of ASH2 Protein Activity

In another embodiment, ASH2 polypeptide levels are determined by virtue of the ASH2 protein activity in a biological sample. Such protein activity can be easily measured using standard techniques. For example, the ASH2 protein has a putative DNA binding domain, *i.e.*, a zinc finger, that likely mediates specific binding to a particular DNA sequence. Thus, ASH2 levels can be indirectly detected in a biological sample by detecting the amount of binding activity to that particular sequence that is present in the sample. Methods of detecting, and quantifying, DNA binding activity are well known to those of skill and are described, *e.g.*, in Sambrook *et al.*, Ausubel *et al.*, *etc.* Other ASH2 activities, *e.g.*, protein binding, can also be assessed using standard methods. In addition, ASH2 is required for the maintenance of gene expression in cells. Accordingly, *in vitro* or *in vivo* transcription based assays (e.g., examining expression of ASH2 target genes) can also be used. Such assays are well known to those of skill in the art.

IV. Diagnosing Cancer

The present invention provides numerous methods for diagnosing any of a number of types of cancer, including, but not limited to, determining whether or not an animal has a cancer, whether or not a biological sample contains cancerous cells, estimating the likelihood of an animal developing cancer, and monitoring the efficacy of anti-cancer treatment in an animal with cancer. Such methods are based on the surprising result that cancer cells have an elevated level of ASH2 polynucleotides (*i.e.*, gene copy number and/or mRNA) and/or polypeptide level. Accordingly, by determining whether or not a cell contains elevated levels of ASH2 polynucleotide or polypeptide, it is possible to determine

whether or not the cell is cancerous. Further, the presence of cancerous cells can be determined indirectly, *i.e.*, a biological sample that does not itself contain cancerous cells, but which has been taken from an animal with cancerous cells elsewhere in its body, may contain elevated levels of ASH2 reflecting the presence of the cancerous cells.

5

A. Detecting a Cancer

In numerous embodiments of the present invention, the level and/or presence or ASH2 polynucleotide or polypeptide is detected in a biological sample, thereby detecting the presence or absence of cancerous cells in the biological sample, or, in certain
10 embodiments, in the animal from which the biological sample was removed. In preferred embodiments, the biological sample comprises a tissue sample from a tissue suspected of containing cancerous cells. For example, in a woman suspected of having breast cancer, breast tissue can be removed. Often, such methods will be used in conjunction with additional diagnostic methods, *e.g.*, detection of other cancer markers, mammography, *etc.*
15 In other embodiments, a tissue sample known to contain cancerous cells, *e.g.*, from a tumor, will be detected for ASH2 levels to determine information about the cancer, *e.g.*, the efficacy of certain treatments, the survival expectancy of the animal, the stage or metastatic potential of the cancer, *etc.*

The amount of ASH2 polynucleotide or polypeptide used to determine the
20 presence of a cancer will depend on numerous factors, including the type of cancer, the age, sex, medical history, *etc.*, of the patient, the cell type, the assay format, *etc.* In preferred embodiments, a level of ASH2 in a biological sample will not be quantified or directly compared with a control sample, but will rather be detected relative to a "diagnostic presence" of ASH2, wherein a "diagnostic presence" refers to the amount of ASH2
25 polynucleotide or polypeptide that indicates the presence of cancer, or the likelihood of cancer, in a particular sample. Preferably, a "diagnostic presence" will be detectable in a simple assay giving a positive or negative result, where a positive "detection" of a "diagnostic presence" of ASH2 polynucleotide or polypeptide indicates the presence of cancer in the animal.

30 The ASH2 level need not be quantified for a "diagnostic presence" to be detected, merely any method of determining whether ASH2 is present at levels higher than in a normal, cancer free cell, sample, or animal. In addition, a "diagnostic presence" does not refer to any absolute quantity of ASH2, but rather on an amount that, depending on the biological sample, cell type, assay conditions, medical condition, *etc.*, is sufficient to

distinguish the level in a cancerous, or pre-cancerous sample, from a normal, cancer-free sample.

Such methods can be practiced regardless of whether any ASH2 polynucleotide or polypeptide is normally present, or "expected" to be present, in a particular control sample. For example, ASH2 may not be expressed in certain cell types, resulting in a complete absence of ASH2 in a control biological sample consisting of such cell types. For such biological sample, a "diagnostic presence" refers to any detectable amount of ASH2. In other tissues, however, there may be a detectable level of ASH2 present in normal, cancer-free cells, and a "diagnostic presence" represents a level that is higher than the normal level, preferably representing a "statistically significant" increase over the normal level. Often, a "diagnostic presence" of ASH2 polynucleotide, polypeptide, and/or protein activity in a biological sample will be at least about 2, 5, 10, or more fold greater than a level expected in a sample taken from a normal, cancer-free animal.

The present methods can also be used to assess the efficacy of a course of treatment. For example, in an animal from which a biological sample has been found to contain an elevated amount of ASH2 polynucleotide or polypeptide, which elevated amount of ASH2 polynucleotide or polypeptide indicated the administration of an anti-cancer therapy to the animal, the efficacy of the treatment can be assessed by monitoring ASH2 levels over time. For example, a reduction in ASH2 polynucleotide or polypeptide levels in a biological sample taken from an animal following a treatment, compared to a level in a sample taken from the animal before the treatment, indicates efficacious treatment.

B. Determining a Prognosis

The level of ASH2 can be used to determine the prognosis of an animal with cancer. For example, if a cancer is detected using a technique other than by detecting ASH2, *e.g.*, tissue biopsy, then the presence or absence of ASH2 can be used to determine the prognosis for the animal with the cancer. For example, an animal with a cancer that has elevated levels of ASH2 may have a reduced survival expectancy compared to an animal with a cancer, but which has normal levels of ASH2, where "survival expectancy" refers to a prediction regarding the severity, duration, or progress of a disease, condition, or any symptom thereof. Methods of correlating ASH2 levels, or the presence or absence of ASH2, with a survival expectancy, likelihood of recurrence of a cancer, and other prognostic factors are well known and can be readily practiced. In some embodiments, alternative

prognostic indicators, *e.g.*, the level or presence of other marker levels, are also detected in conjunction with the detection of ASH2.

C. Determining a Preferred Course of Treatment

The present methods can be used to determine the optimal course of treatment in an animal with cancer. For example, the detection of an elevated level of ASH2 can indicate a reduced survival expectancy of an animal with a cancer, thereby indicating a more aggressive treatment for the animal. In addition, a correlation can readily be established between levels of ASH2, or the presence or absence of an elevated level of ASH2, and the relative efficacy of one or another anti-cancer agent. Such analyses can be performed, *e.g.*, retrospectively, *i.e.*, by detecting ASH2 levels in samples taken previously from animals that have subsequently undergone one or more types of anti-cancer therapy, and correlating the ASH2 levels with the known efficacy of the treatment.

V. Treating cancer

The present invention provides numerous methods for treating an animal with a cancer. In addition to allowing the determination of an optimal treatment for an animal with cancer, as described *supra*, methods are provided for treating a cancer by inhibiting the growth and/or proliferation of a cancer cell. As used herein, such methods are directed at reducing the level of ASH2 polypeptide levels, polynucleotide levels, or protein activity in a cancerous cell. It will be appreciated that more than one of the methods described *infra* can be performed on a given animal, and may also be administered in conjunction with one or more traditional, well known anti-cancer therapies, *e.g.*, chemotherapy, radiation therapy, surgery, hormone therapy, immunotherapy, *etc.*

According to the present invention, a "method of treating cancer" refers to a procedure or course of action that is designed to reduce or eliminate the number of cancer cells in an animal, or to alleviate the symptoms of a cancer. "A method of treating cancer" does not necessarily mean that the cancer cells will, in fact, be eliminated, that the number of cells will, in fact, be reduced, or that the symptoms of a cancer will, in fact, be alleviated. Often, a method of treating cancer will be performed even with a low likelihood of success, but which, given the medical history and estimated survival expectancy of an animal, is deemed an overall beneficial course of action.

In certain embodiments, the present invention provides methods for treating cancer by detecting the level and/or a diagnostic presence of ASH2 polynucleotide or

polypeptide in a biological sample, and, when a diagnostic presence is detected, administering one or more of the above-listed anti-cancer therapies.

One commonly applied anti-cancer therapy is chemotherapy. As used herein, "chemotherapy" refers to the administration of chemical compounds to an animal with cancer that is aimed at killing or reducing the number of cancer cells within the animal. Generally, chemotherapeutic agents arrest the growth of or kill cells that are dividing or growing, such as cancer cells. Examples of chemotherapeutic agents include doxirubicin, vinblastine, genistein, taxol, vincristine, *etc.*

Another commonly applied anti-cancer therapy is radiation therapy. "Radiation therapy" refers to the administration of radioactivity to an animal with cancer. Radiation kills or inhibits the growth of dividing cells, such as cancer cells. The administration of radiation may be from an external source (*e.g.*, a gamma source, a proton source, a molecular beam source, *etc.*), or may be through an implantable radioactive material.

In numerous embodiments, a tissue found to be cancerous using the present methods will be removed using surgery. "Surgery" refers to the direct removal or ablation of cells, *e.g.*, cancer cells, from an animal. Most often, the cancer cells will be in the form of a tumor (*e.g.*, a mammary tumor), which is removed from the animal. The surgical methods may involve removal of healthy as well as cancerous tissue.

Hormone therapy can also be used to treat cancers, *e.g.*, breast cancer. As used herein, "hormone therapy" refers to the administration of compounds that counteract or inhibit hormones, such as estrogen or androgen, that have a mitogenic effect on cells. Often, such hormones act to increase the cancerous properties of cancer cells *in vivo*. "Hormone therapy" can also include methods of reducing or eliminating the production of hormones in an animal, *e.g.*, the surgical removal of ovaries to prevent estrogen production.

In certain embodiments, immunotherapy will be used to treat cancer, *e.g.*, a cancer detected using the present methods. "Immunotherapy" refers to methods of enhancing the ability of an animal's immune system to destroy cancer cells within the animal. Numerous such methods are well known to those of skill in the art. This can involve the treatment with polyclonal or monoclonal antibodies (*e.g.*, Herceptin) that bind to particular molecules located on, produced by, or indicative of, tumor cells. *See, e.g.*, Pastan *et al.*(1992) *Ann. Rev. Biochem.*, 61: 331-354, Brinkman and Pastan (1994) *Biochimica Biphysica Acta,* 1198: 27-45, *etc.*

A. Reducing ASH2 Levels in Cells

In preferred embodiments, this invention provides methods of treating a cancer by reducing ASH2 levels in a cell. Typically, such methods are used to reduce an elevated level of ASH2, *e.g.*, an elevated level in a cancerous cell. According to the present invention, "reducing the level of ASH2 activity" refers to inhibiting ASH2 protein activity in the cell, lowering the copy number of ASH2 genes, or decreasing the level of ASH2 mRNA or protein in the cell. Preferably, the level of ASH2 activity is lowered to the level typical of a normal, cancer-free cell, but the level may be reduced to any level that is sufficient to decrease the proliferation of the cell, including in certain cases to levels below those typical of normal cells. Preferably, such methods involve the use of inhibitors of ASH2, where an "inhibitor of ASH2" is a molecule that acts to reduce ASH2 polynucleotide levels, ASH2 polypeptide levels and/or ASH2 protein activity. Such an inhibitor can include antisense polynucleotides, ribozymes, antibodies, dominant negative ASH2 forms, and small molecule inhibitors of ASH2.

In preferred embodiments, ASH2 levels will be reduced so as to reduce the proliferation of a cancer cell with elevated ASH2 levels. The "proliferation" of a cell refers to the rate at which the cell or population of cells grows and divides, or to the extent to which the cell or population of cells grows, divides or increases in number. "Proliferation" can reflect multiple factors, including the rate of cell growth and division and the rate of cell death. A method of "decreasing" the proliferation of a cell means to reduce the rate or extent of growth or division of a cell or population of cells. Such methods can involve preventing cell division or cell growth, and may also include cell killing, and can be practiced *in vivo* or *in vitro*.

In preferred embodiments, ASH2 levels will be reduced in a tumor cell, a hyperproliferative cell, and/or a metastatic cell. A "tumor cell" is a cancer cell, *in vitro* or *in vivo*, that is part of a tumor, has been isolated from a tumor, or which is capable of forming a tumor. A "hyperproliferative cell" is a cell with an abnormally high rate of proliferation, or a cell that proliferates to an abnormally great extent, *i.e.*, gives rise to a population of cells that increases in number over time. Typically, cancer cells are metastatic, *i.e.*, capable of leaving their normal anatomical location and moving to, and proliferating in, another part of an animal. Typically, such "metastatic" cells have acquired the ability to cross basal laminae so as to leave their normal tissue, enter the circulation, leave the circulation, and proliferate in a new location.

1. Antisense Polynucleotides

In certain embodiments, ASH2 activity is downregulated, or entirely inhibited, by the use of antisense polynucleotide. An "antisense polynucleotide" is a nucleic acid complementary to, and which can preferably hybridize specifically to, a coding mRNA nucleic acid sequence, *e.g*, ASH2 mRNA, or a subsequence thereof. Binding of the antisense polynucleotide to the ASH2 mRNA reduces the translation and/or stability of the ASH2 mRNA.

In the context of this invention, antisense polynucleotides can comprise naturally-occurring nucleotides, or synthetic species formed from naturally-occurring subunits or their close homologs. Antisense polynucleotides may also have altered sugar moieties or inter-sugar linkages. Exemplary among these are the phosphorothioate and other sulfur containing species which are known for use in the art. All such analogs are comprehended by this invention so long as they function effectively to hybridize with ASH2 mRNA.

Such antisense polynucleotides can readily be synthesized using recombinant means, or can be synthesized *in vitro*. Equipment for such synthesis is sold by several vendors, including Applied Biosystems. The preparation of other oligonucleotides such as phosphorothioates and alkylated derivatives is also well known to those of skill in the art.

2. Ribozymes

In addition to antisense polynucleotides, ribozymes can be used to target and inhibit transcription of ASH2. A ribozyme is an RNA molecule that catalytically cleaves other RNA molecules. Different kinds of ribozymes have been described, including group I ribozymes, hammerhead ribozymes, hairpin ribozymes, RNAse P, and axhead ribozymes (*see* Castanotto *et al.* (1994) *Adv. in Pharmacology* 25: 289-317 for a general review of the properties of different ribozymes).

The general features of hairpin ribozymes are described, *e.g.*, in Hampel *et al.* (1990) *Nucl. Acids Res.* 18: 299-304; Hampel *et al.* (1990) European Patent Publication No. 0 360 257; U.S. Patent No. 5,254,678. Methods of preparing ribozymes are well known to those of skill in the art (*see, e.g.,* Wong-Staal *et al.*, WO 94/26877; Ojwang *et al.* (1993) *Proc. Natl. Acad. Sci.* USA 90: 6340-6344; Yamada *et al.* (1994) *Human Gene Therapy* 1: 39-45; Leavitt *et al.* (1995) *Proc. Natl. Acad. Sci. USA* 92: 699-703; Leavitt *et al.* (1994) *Human Gene Therapy* 5: 1151-120; and Yamada *et al.* (1994) *Virology* 205: 121-126).

3. Inhibitors of ASH2 Polypeptide Activity

ASH2 activity can also be decreased by the addition of an inhibitor of the ASH2 polypeptide. This can be accomplished in any of a number of ways, including by providing a dominant negative ASH2 polypeptide, *e.g.*, a form of ASH2 that itself has no activity and which, when present in the same cell as a functional ASH2, reduces or eliminates the ASH2 activity of the functional ASH2. Also, inactive polypeptide variants (muteins) can be used, *e.g.*, by screening for the ability to inhibit ASH2 activity. Methods of making muteins are well known to those of skill (*see, e.g.,* U.S. Patent Nos. 5,486,463, 5,422,260, 5,116,943, 4,752,585, 4,518,504).

a) Screening for ASH2 Inhibitors

In one embodiment, this invention provides methods of screening for agents that modulate and preferably downregulate ASH2 protein activity or expression. Preferred "screening" methods of this invention involve (i) contacting an ASH2-expressing cell (*e.g.,* a cell capable of expressing ASH2) with a test agent; and (ii) detecting the level of ASH2 activity (*e.g.,* as described above), where a decreased level of ASH2 activity as compared to the level of ASH2 activity in a cell not contacted with the test agent indicates that the test agent inhibits or downregulates ASH2.

Virtually any agent can be tested in such an assay. Such agents include, but are not limited to, natural or synthetic nucleic acids, natural or synthetic polypeptides, natural or synthetic lipids, natural or synthetic small organic molecules, and the like. In one preferred format, test agents are provided as members of a combinatorial library. In preferred embodiments, a collection of small molecule inhibitors are tested for ASH2 inhibiting ability. A "small molecule inhibitor" of ASH2 is any molecule, *e.g.*, a carbohydrate, nucleotide, amino acid, oligonucleotide, oligopeptide, lipid, inorganic compound, *etc.* that inhibits ASH2 protein activity. Such molecules can inhibit ASH2 protein activity by any of a number of mechanisms, *e.g.*, by binding to an ASH2 protein and competitely inhibiting its interaction with DNA or with other proteins. Preferably, such "small molecule inhibitors" are smaller than about 10 kD. More preferably, such inhibitors are smaller than about 5, 2, or 1 kD or even smaller.

Test agents can also be screened based on functional properties of ASH2 protein. For example, loss of ASH2 function in the fruitfly *Drosophila melanogaster* results in "homeotic transformations," *i.e.*, alterations in the identity of one or more parts, *i.e.*, segments, of the Drosophila body plan. Accordingly, molecules that reduce or eliminate

ASH2 function in Drosophila will reproduce these same alterations, even in the presence of wild type ASH2 function. In such embodiments, a candidate molecule or agent will be administered to, *e.g.*, a Drosophila embryo (*e.g.*, applying a compound to an embryo, or expressing a candidate molecule in cells of a Drosophila embryo), and the appearance of one or more homeotic transformations scored. Any treatment that results in homeotic transformations similar to those seen in ASH2 mutants can be further analyzed for their ability to directly inhibit ASH2 expression or activity.

(i) Combinatorial Libraries

In certain embodiments, combinatorial libraries of potential ASH2 modulators will be screened for ASH2-inhibiting ability. Conventionally, new chemical entities with useful properties are generated by identifying a chemical compound (called a "lead compound") with some desirable property or activity, *e.g.*, ASH2 inhibiting activity, creating variants of the lead compound, and evaluating the property and activity of those variant compounds. However, the current trend is to shorten the time scale for all aspects of drug discovery. Because of the ability to test large numbers quickly and efficiently, high throughput screening (HTS) methods are replacing conventional lead compound identification methods.

In one preferred embodiment, high throughput screening methods involve providing a library containing a large number of potential therapeutic compounds (candidate compounds). Such "combinatorial chemical libraries" are then screened in one or more assays to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional "lead compounds" or can themselves be used as potential or actual therapeutics.

A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis by combining a number of chemical "building blocks" such as reagents. For example, a linear combinatorial chemical library, such as a polypeptide (*e.g.*, mutein) library, is formed by combining a set of chemical building blocks called amino acids in every possible way for a given compound length (*i.e.*, the number of amino acids in a polypeptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks (Gallop *et al.* (1994) *J. Med. Chem.* 37(9): 1233-1251).

Preparation and screening of combinatorial chemical libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to, peptide libraries (*see, e.g.*, U.S. Patent No. 5,010,175, Furka (1991) *Int. J. Pept. Prot. Res.*, 37: 487-493, Houghton *et al.* (1991) *Nature*, 354: 84-88), peptoids (PCT Publication No WO 91/19735, 26 Dec. 1991), encoded peptides (PCT Publication WO 93/20242, 14 Oct. 1993), random bio-oligomers (PCT Publication WO 92/00091, 9 Jan. 1992), benzodiazepines (U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs *et al.*, (1993) *Proc. Nat. Acad. Sci. USA* 90: 6909-6913), vinylogous polypeptides (Hagihara *et al.* (1992) *J. Amer. Chem. Soc.* 114: 6568), nonpeptidal peptidomimetics with a Beta-D-Glucose scaffolding (Hirschmann *et al.*, (1992) *J. Amer. Chem. Soc.* 114: 9217-9218), analogous organic syntheses of small compound libraries (Chen *et al.* (1994) *J. Amer. Chem.* Soc. 116: 2661), oligocarbamates (Cho, et al., (1993) *Science* 261:1303), and/or peptidyl phosphonates (Campbell *et al.*, (1994) *J. Org. Chem.* 59: 658). *See, generally,* Gordon *et al.*, (1994) *J. Med. Chem.* 37:1385, nucleic acid libraries (*see, e.g.*, Strategene, Corp.), peptide nucleic acid libraries (*see, e.g.*, U.S. Patent 5,539,083), antibody libraries (*see, e.g.*, Vaughn *et al.* (1996) *Nature Biotechnology*, 14(3): 309-314), and PCT/US96/10287), carbohydrate libraries (*see, e.g.*, Liang *et al.*, (1996) *Science*, 274: 1520-1522, and U.S. Patent No. 5,593,853), and small organic molecule libraries (*see, e.g.*, benzodiazepines, Baum (1993) C&EN, Jan 18, page 33; isoprenoids, U.S. Patent No. 5,569,588; thiazolidinones and metathiazanones, U.S. Patent No. 5,549,974; pyrrolidines, U.S. Patent Nos. 5,525,735 and 5,519,134; morpholino compounds, U.S. Patent No. 5,506,337; benzodiazepines, U.S. Patent No. 5,288,514; and the like).

Devices for the preparation of combinatorial libraries are commercially available (*see, e.g.*, 357 MPS, 390 MPS, Advanced Chem Tech, Louisville KY, Symphony, Rainin, Woburn, MA, 433A Applied Biosystems, Foster City, CA, 9050 Plus, Millipore, Bedford, MA).

A number of well known robotic systems have also been developed for solution phase chemistries. These systems include automated workstations like the automated synthesis apparatus developed by Takeda Chemical Industries, LTD. (Osaka, Japan) and many robotic systems utilizing robotic arms (Zymate II, Zymark Corporation, Hopkinton, Mass.; Orca, Hewlett-Packard, Palo Alto, Calif.), which mimic the manual synthetic operations performed by a chemist. Any of the above devices are suitable for use with the present invention. The nature and implementation of modifications to these devices

(if any) so that they can operate as discussed herein will be apparent to persons skilled in the relevant art. In addition, numerous combinatorial libraries are themselves commercially available (*see, e.g.,* ComGenex, Princeton, N.J., Asinex, Moscow, Ru, Tripos, Inc., St. Louis, MO, ChemStar, Ltd, Moscow, RU, 3D Pharmaceuticals, Exton, PA, Martek Biosciences, Columbia, MD, *etc.*).

(ii) High Throughput Screening

Any of the assays to identify compounds capable of modulating ASH2 levels described herein are amenable to high throughput screening. Preferred assays thus detect enhancement or inhibition of ASH2 gene transcription, inhibition or enhancement of ASH2 polypeptide expression, and inhibition or enhancement of ASH2 polypeptide activity.

High throughput assays for the presence, absence, quantification, or other properties of particular nucleic acids or protein products are well known to those of skill in the art. Similarly, binding assays and reporter gene assays are similarly well known. Thus, for example, U.S. Patent No. 5,559,410 discloses high throughput screening methods for proteins, U.S. Patent No. 5,585,639 discloses high throughput screening methods for nucleic acid binding (*i.e.,* in arrays), while U.S. Patent Nos. 5,576,220 and 5,541,061 disclose high throughput methods of screening for ligand/antibody binding.

In addition, high throughput screening systems are commercially available (*see, e.g.,* Zymark Corp., Hopkinton, MA; Air Technical Industries, Mentor, OH; Beckman Instruments, Inc. Fullerton, CA; Precision Systems, Inc., Natick, MA, *etc.*). These systems typically automate entire procedures, including all sample and reagent pipetting, liquid dispensing, timed incubations, and final readings of the microplate in detector(s) appropriate for the assay. These configurable systems provide high throughput and rapid start up as well as a high degree of flexibility and customization. The manufacturers of such systems provide detailed protocols for various high throughput systems. Thus, for example, Zymark Corp. provides technical bulletins describing screening systems for detecting the modulation of gene transcription, ligand binding, and the like.

B. Administration of ASH2-inhibiting Compounds

In numerous embodiments of the present invention, an ASH2 inhibiting compound, *i.e.,* a compound that reduces levels of ASH2 mRNA, polypeptide and/or protein activity, will be administered to an animal. Such compounds can be administered by a variety of methods including, but not limited to, parenteral, topical, oral, or local

administration, such as by aerosol or transdermally, for prophylactic and/or therapeutic treatment. The pharmaceutical compositions can be administered in a variety of unit dosage forms depending upon the method of administration. For example, unit dosage forms suitable for oral administration include, but are not limited to, powder, tablets, pills, capsules and lozenges. It is recognized that the ASH2 modulators (*e.g.*, antibodies, antisense constructs, ribozymes, small organic molecules, *etc.*) when administered orally, must be protected from digestion. This is typically accomplished either by complexing the molecule(s) with a composition to render it resistant to acidic and enzymatic hydrolysis, or by packaging the molecule(s) in an appropriately resistant carrier, such as a liposome. Means of protecting agents from digestion are well known in the art.

The compositions for administration will commonly comprise an ASH2 modulator dissolved in a pharmaceutically acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers can be used, *e.g.*, buffered saline and the like. These solutions are sterile and generally free of undesirable matter. These compositions may be sterilized by conventional, well known sterilization techniques. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions such as pH adjusting and buffering agents, toxicity adjusting agents and the like, for example, sodium acetate, sodium chloride, potassium chloride, calcium chloride, sodium lactate and the like. The concentration of active agent in these formulations can vary widely, and will be selected primarily based on fluid volumes, viscosities, body weight and the like in accordance with the particular mode of administration selected and the patient's needs.

A typical pharmaceutical composition for intravenous administration would be about 0.1 to 10 mg per patient per day. Dosages from 0.1 up to about 100 mg per patient per day may be used, particularly when the drug is administered to a secluded site and not into the blood stream, such as into a body cavity or into a lumen of an organ. Substantially higher dosages are possible in topical administration. Actual methods for preparing parenterally administrable compositions will be known or apparent to those skilled in the art and are described in more detail in such publications as *Remington's Pharmaceutical Science*, 15th ed., Mack Publishing Company, Easton, Pennsylvania (1980).

The compositions containing modulators of ASH2 can be administered for therapeutic or prophylactic treatments. In therapeutic applications, compositions are administered to a patient suffering from a disease (*e.g.*, a cancer) in an amount sufficient to cure or at least partially arrest the disease and its complications. An amount adequate to accomplish this is defined as a "therapeutically effective dose." Amounts effective for this

use will depend upon the severity of the disease and the general state of the patient's health. Single or multiple administrations of the compositions may be administered depending on the dosage and frequency as required and tolerated by the patient. In any event, the composition should provide a sufficient quantity of the agents of this invention to effectively treat the patient.

VI. Cancers

The present methods can be used to diagnose and treat any of a number of types of cancers. In preferred embodiments, epithelial cancers will be diagnosed and/or treated. For example, breast, ovarian, colorectal, kidney, stomach, bladder, lung, and any other epithelial cancer can be treated or diagnosed. In addition, a cancer at any stage of progression can be detected, such as primary, metastatic, and recurrent cancers. Information regarding numerous types of cancer can be found, *e.g.*, from the American Cancer Society (www3.cancer.org), or from, *e.g.*, Wilson *et al.* (1991) Harrison's Principles of Internal Medicine, 12th Edition, McGraw-Hill, Inc.

VII. Novel ASH2 Polynucleotides and Polypeptides

In addition to the foregoing, the present invention provides novel variants of ASH2. The ASH2 polynucleotides and polypeptides provided herein have numerous uses, including for diagnosing cancer, determining the prognosis of an animal with cancer, determining the efficacy of a cancer treatment, inhibiting the proliferation of a cell *in vitro* or *in vivo*, and others. In addition, the present ASH2 polynucleotides and polypeptides can be synthesized and purified using standard techniques (*see, e.g.*, Sambrook, Ausubel, both *supra*). One can also use the polypeptides of the invention to prepare antibodies that recognize ASH2; such antibodies find use in immunoassays, for example, to detect ASH2 in a biological sample.

It has been discovered that ASH2 comprises a number of splice variants, any of which can be used in the present invention. The novel ASH2 variants provided herein were identified, *inter alia*, as novel cDNA clones isolated from a human peripheral blood leukocyte cDNA library. For example, the cDNA shown as SEQ ID NO:1, which comprises 1627 nucleotides and encodes a protein sequence of 305 amino acids (shown as SEQ ID NO:2), is similar to several previously described human ASH2 sequences (*e.g.*, GenBank Accession Nos: AF056717, AB020982, and AF056718) that were of unknown function.

This sequences, and others provided herein, differ from the previously identified sequences, however, in several important and unanticipated respects, as described *infra*.

First, certain of the ASH2 polynucleotides provided herein comprise a different translation start site than most previously described ASH2 sequences. For example, the human ASH2 polypeptide sequence shown as SEQ ID NO:2 lacks the first 94 amino acids of other ASH2 sequences, *e.g.*, GenBank Accession No. AB022785. In addition, all previously described ASH2 sequences contain an exon, shown as SEQ ID NO:4 and representing, *e.g.*, amino acids 389-510 in AB022785, that is not present in the sequences provided herein. In addition, the terminal 11 amino acids of the ASH2 polypeptide shown as SEQ ID NO:2 (the 11 amino acid sequence is shown as SEQ ID NO:3) comprises a novel ASH2 subsequence not previously described in any protein. Finally, the subsequence present in SEQ ID NO:8, shown alone as SEQ ID NO:9, also comprises a novel ASH2 subsequence never before described.

The ASH2 protein has a putative double zinc-finger domain, called a PHD finger (*see, e.g.*, Adamson *et al.*, *Genetics* 1996 Oct;144(2):621-33. Zinc fingers are generally thought to mediate DNA binding, suggesting that the ASH2 protein binds to DNA, *e.g.*, to regulate the expression of one or more target genes. It will be appreciated that the ASH2 polypeptide or polynucleotides sequences provided herein can represent fragments of a full length ASH2. Typically, such fragments will represent one or more discrete subdomains of ASH2, *e.g.*, one or both zinc finger domains can be used independently of the entire protein.

The manipulation of any of the polynucleotides or proteins described herein, *e.g.*, for the purpose of producing protein or nucleic acids, or for creating variants, derivatives, fragments, *etc.*, can be accomplished using standard molecular biological techniques, as described, *e.g.*, in Ausubel *et al.* (ed.) (1990) Current Protocols in Molecular biology, Greene Publishing and Wiley-Interscience, New York, Glover (ed.) (1987) *DNA Cloning: A Practical Approach,* vols 1-3, IRL Press, Oxford, or Sambrook *et al.* (1989) Molecular Cloning: A Laboratory Manual, 2d Ed., vols 1-3, Cold Spring Harbor Press, New York. The ASH2 polynucleotides and polypeptides of this invention include isolated ASH2 polynucleotides and polypeptides, wherein "isolated" indicates that the polynucleotides or polypeptides are substantially free of other polypeptides and other cellular components with which they are naturally associated.

1. Introducing Nucleic Acids into Cells

In numerous embodiments, one or more nucleic acids, *e.g.*, ASH2 polynucleotides, such as antisense polynucleotides or ribozymes, will be introduced into cells, *in vitro* or *in vivo*. The present invention provides methods, reagents, vectors, and cells useful for expression of ASH2 and other polypeptides and nucleic acids using *in vitro* (cell-free), *ex vivo* or *in vivo* (cell or organism-based) recombinant expression systems.

The particular procedure used to introduce the nucleic acids into a host cell for expression of a protein or nucleic acid is application specific. Any of the well known procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, spheroplasts, electroporation, liposomes, microinjection, plasma vectors, viral vectors and any of the other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA or other foreign genetic material into a host cell (*see, e.g.*, Berger and Kimmel, *Guide to Molecular Cloning Techniques, Methods in Enzymology* volume 152 Academic Press, Inc., San Diego, CA (Berger), F.M. Ausubel *et al.*, eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (supplemented through 1999), and Sambrook *et al., Molecular Cloning - A Laboratory Manual* (2nd Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989.

a) Vectors

In numerous embodiments of this invention, nucleic acids encoding ASH2 polypeptides, or inhibitors thereof, will be inserted into vectors using standard molecular biological techniques. Vectors may be used at multiple stages of the practice of the invention, *e.g.*, for subcloning nucleic acids encoding ASH2 polypeptides, ,or ASH2 inhibitors, *e.g.*, ASH2 ribozymes or antisense sequences, or for subcloning additional elements used to control protein or mRNA expression, vector selectability, *etc.* Vectors may also be used to maintain or amplify the nucleic acids, for example, by inserting the vector into prokaryotic or eukaryotic cells and growing the cells in culture. In addition, vectors may be used to introduce and express ASH2 nucleic acids, or ASH2-inhibiting nucleic acids, *e.g.*, ASH2 ribozymes or antisense sequences, into cells for therapeutic or experimental purposes.

A variety of commercially or commonly available vectors and vector nucleic acids can be converted into a vector of the invention by cloning a polynucleotide of this invention into the commercially or commonly available vector. A variety of common

vectors suitable for this purpose are well known in the art. For cloning in bacteria, common vectors include pBR322-derived vectors such as pBLUESCRIPT™, and bacteriophage derived vectors. In yeast, vectors include Yeast Integrating plasmids (*e.g.*, YIp5) and Yeast Replicating plasmids (the YRp series plasmids) and pGPD-2. Expression in mammalian cells can be achieved using a variety of commonly available plasmids, including pSV2, pBC12BI, and p91023, as well as lytic virus vectors (*e.g.*, vaccinia virus, adeno virus, and baculovirus), episomal virus vectors (*e.g.*, bovine papillomavirus), and retroviral vectors (*e.g.*, murine retroviruses).

Typically, a nucleic acid subsequence encoding an ASH2 polypeptide is placed under the control of a promoter. A nucleic acid is "operably linked" to a promoter when it is placed into a functional relationship with the promoter. For instance, a promoter or enhancer is operably linked to a coding sequence if it increases or otherwise regulates the transcription of the coding sequence. Similarly, a "recombinant expression cassette" or simply an "expression cassette" is a nucleic acid construct, generated recombinantly or synthetically, with nucleic acid elements that are capable of effecting expression of a structural gene in hosts compatible with such sequences. Expression cassettes include promoters and, optionally, introns, polyadenylation signals, and transcription termination signals. Additional factors necessary or helpful in effecting expression may also be used as described herein. For example, an expression cassette can also include nucleotide sequences that encode a signal sequence that directs secretion of an expressed protein from the host cell. Transcription termination signals, enhancers, and other nucleic acid sequences that influence gene expression, can also be included in an expression cassette.

An extremely wide variety of promoters are well known, and can be used in the vectors of the invention, depending on the particular application. Ordinarily, the promoter selected depends upon the cell in which the promoter is to be active. Other expression control sequences such as ribosome binding sites, transcription termination sites and the like are also optionally included. For *E. coli*, example control sequences include the T7, trp, or lambda promoters, a ribosome binding site and preferably a transcription termination signal. For eukaryotic cells, the control sequences typically include a promoter which optionally includes an enhancer derived from immunoglobulin genes, SV40, cytomegalovirus, a retrovirus (*e.g.*, an LTR based promoter) etc., and a polyadenylation sequence, and may include splice donor and acceptor sequences.

VIII. Kits for Use in Diagnostic and/or Prognostic Applications

For use in diagnostic, research, and therapeutic applications suggested above, kits are also provided by the invention. In the diagnostic and research applications such kits may include any or all of the following: assay reagents, buffers, ASH2 specific nucleic acids or antibodies, hybridization probes and/or primers, antisense polynucleotides, ribozymes, dominant negative ASH2 polypeptides or polynucleotides, small molecules inhibitors of ASH2, *etc.* A therapeutic product may include sterile saline or another pharmaceutically acceptable emulsion and suspension base.

In addition, the kits may include instructional materials containing directions (*i.e.,* protocols) for the practice of the methods of this invention. While the instructional materials typically comprise written or printed materials they are not limited to such. Any medium capable of storing such instructions and communicating them to an end user is contemplated by this invention. Such media include, but are not limited to electronic storage media (*e.g.,* magnetic discs, tapes, cartridges, chips), optical media (*e.g.*, CD ROM), and the like. Such media may include addresses to internet sites that provide such instructional materials.

IX. EXAMPLES

A. Discovery of Amplified Region

By representational difference analysis (RDA) of a human breast cancer biopsy, an RDA probe that detects amplification in several breast cancer samples has been discovered.

Construction of a physical map of this region revealed *FGFR1* is ~0.6 Mb away from the RDA probe (Figure 1). There are two types of experimental analysis that indicate that *FGFR1* is not the only oncogene that is driving amplification of this region (if at all). For one, analysis of copy-number levels by quantitative PCR (Q-PCR) with fluorogenic TaqMan probes indicated that there are sequences in between the RDA probe and *FGFR1* that are not amplified or less amplified than the RDA probe and *FGFR1*. This suggests that the amplification of the two regions is being driven by two separate oncogenes. Second, one out of two tumors (88-688 and 89-249) analyzed by quantitative RT-PCR did not have the *FGFR1* gene overexpressed (89-249). This lack of overexpression rules out *FGFR1* as the oncogene responsible for amplification of the region in this tumor sample.

Figure 1 shows a physical map of the 8p11-p12 amplified chromosomal region, indicated by the solid line marked with solid circles. Underneath are solid lines representing individual YACs, and a dashed line respresenting genomic sequence from GenBank (Accession numbers: AP000065 through AP000084). STSs on this map (solid circles) were derived from RDA probes or from the genomic sequence or public databases at the Whitehead Institute, Stanford University, and the Sanger Center. The STSs are ordered on the basis of genomic sequence; or by determination by PCR of their presence on these YACs and others not shown, and confirmed by radiation hybrid mapping using the GB3 panel (Stanford University RH server).

92 cell lines were scanned for copy-number increases using Q-PCR and fluorogenic TaqMan probes to the RDA probe-derived STS, RDA-534. Two out of 26 breast cancer cell lines (BT-483 and MDA-MB-134), 1 out of 15 colorectal cancer cell lines (SW837), and 1 out of 32 lung cancer cell lines (SK-LC14) were amplified greater than three-fold. 220 primary tumors were also surveyed for copy number increases using a TaqMan probe derived from the sequence of STS NIB1979 (3' UTR of ASH2). 21 of the 220 tumors were amplified 2.5-fold or greater, and of these 21, 10 were amplified five-fold or greater. TaqMan probes (ABI, Foster City, Calif.) for the other 12 STSs were used to measure additional copy-number increases in six amplified primary tumors, and the results of these analyses are noted in Figure 1 (-Fold Amplification).

B. Significant Overexpression of ASH2 in Amplified Tumors

The ASH2 gene was initially identified as amplified in tumors by searching databases for gene identities of the public ESTs that mapped near to the RDA probe. The public ESTs that were examined had been previously mapped nearby at low resolution (as part of the HGP gene-mapping effort). The EST NIB1979 mapped nearby RDA probe 534 and is derived from the 3' UTR of ASH2.

Northern analyses were performed to determine the expression level of ASH2. A human-tissue panel, as well as a panel of 26 breast cancer cell lines, including normal human mammary epithelial cells for normalization of RNA expression levels, were used in these experiments. ASH2 was found to be expressed as a ~ 3.0 kb transcript at similar levels in all human tissues examined (pancreas, kidney, muscle, liver, lung, placenta, brain, and heart). ASH2 was significantly overexpressed in all cancer cell lines in which it is amplified, as well as in three breast cancer cell lines in the absence of gene amplification

(Table 1). On the basis of this analysis, it was determined that ASH2 is overexpressed in ~25% of breast cancer cell lines.

Table 1. Northern analysis of expression levels of ASH2 in cancer cell lines

Cell line	Tumor type	ASH2 amplification	ASH2 Expression
HMEC	Normal Breast	1x	1x
BT483	Breast	6x	24x
MDAMB134	Breast	11x	24x
SW837	Colorectal	4x	9x
CAMA-1	Breast	2.5x	16x
DU4475	Breast	1x	8x
MDAMB157	Breast	1x	10x
MDAMB330	Breast	1x	10x

Northern analysis of total RNA was quantitated with a Fuji phosphorimager, and the expression level is the ratio of the test gene to GAPDH normalized to human mammary epithelial cell (HMEC) levels.

Quantitative RT-PCR analyses were also performed to determine the expression level *of ASH2* in human primary tumor biopsies. This analysis indicated that ASH2 is overexpressed in those tumors in which it is amplified (Table 2).

Table 2. TaqMan quantitative RT-PCR analysis of expression levels of ASH2 in human breast cancer biopsies

Tumor or tissue sample	ASH2 amplification	ASH2 Expression
Normal Mammary	1x	1x
88-688	11x	12x
89-249	5x	3x
96-76	7x	11x
96-194	1x	0.3x
96-360	1x	1x

TaqMan RT-PCR analysis of total RNA was performed according to protocols from Perkin Elmer, Foster City, CA. Expression level is the ratio of the ASH2 gene to GAPDH normalized to human mammary gland levels (Clontech, Palo Alto, CA).

C. Structure and Sequence of the ASH2 Gene

To date numerous novel ASH2 cDNA clones have been sequenced, representing several novel ASH2 isoforms. One of these, isolated from a human peripheral blood leukocyte cDNA library, is shown as SEQ ID NO:1, and is similar to submitted sequences AF056717, AB020982, and AF056718 (all human sequences homologous to Drosophila ASH2). It is believed that this sequence, which is a hybrid of AF056717 and AB020982 yet is missing putative exons found in both, is an alternate splicing isoform that is found leukocytes.

Two additional novel isoforms of ASH2, referred to as SE 707 and SE 691, have also been identified. SE 707, shown as SEQ ID NO:5 (cDNA) and SEQ ID NO:6 (amino acid sequence), is identical to AB020982 (*see*, GenBank) except that it skips exons 11 and 12, corresponding to nucleotides 1177-1538 0f AB020982. This alternate splice produces a frameshift, resulting in a termination codon (the stop codon is shown as an asterix in the amino acid SEQ ID NO:6) 11 amino acids downstream of the splice. The nucleotides at the novel splice junction are underlined in SEQ ID NO:5, as is the novel C-terminal peptide in SEQ ID NO:6 (the novel 11 amino acid peptide is shown separately as SEQ ID NO:3).

Isoform SE 691 (the cDNA of which is shown as SEQ ID NO:7 and the amino acid sequence as SEQ ID NO:8) is identical to SE 707, except that it has an insertion of 109 nucleotides between exons 1 and 2 (between nucleotides 199 and 200 of AB020982), resulting in a frameshift and a premature termination codon. This nucleotide insertion, and the resulting novel polypeptide sequence, are underlined in SEQ ID NO:7 and SEQ ID NO:8, respectively. The novel polypeptide sequence is also shown separately as SEQ ID NO:9.

While the foregoing invention has been described in some detail for purposes of clarity and understanding, it will be clear to one skilled in the art from a reading of this disclosure that various changes in form and detail can be made without departing from the true scope of the invention. For example, all the techniques and apparatus described above may be used in various combinations. All publications and patent documents cited in this application are incorporated by reference in their entirety for all purposes to the same extent as if each individual publication or patent document were so individually denoted.

SEQ ID NO:1

ASH2 cDNA clone

ATAATTCGGCACGAGCCGTGATGGCGGCGGCAGGAGCAGGACCTGGCCAGGAA
GCGGGTGCCGGGCCTGGCCCAGGAGCGGTCGCAAATGCAACAGGGGCAGAAGA
GGGGGAGATGAAGCCGGTGGCAGCGGGAGCAGCCGCTCCTCCTGGAGAGGGGA
TNTNTGCTGCTCCGACAGTTGAGCCCAGTTCCGGGGAGGCTGAAGGCGGAAAAG
AAACCACTNNGCAGTGCCAGACTGGAAGAAGTAACGGTCACTCTGAAAACAGG
GTGGGAGAGCTGCCTCTCTTTGAACNTCTCCCAGGACCAAYTCTAACCCNGGGA
GGCAAACTTGGTNGANGTAAGCGGTGGCTTGGAGACAGAATCATTTAATGGAA
AAGATACACTAGAAGGTGCTGGGGATACATCAGAGGTGATGGATACTCAGGCG
GGNTCCGTGGATGAAGAGAATGGCCGACAGTTGGGTGAGGTAGAGCTGCAATG
TGGGATTTGTACAAAATGGTTCACGGCTGACACATTTGGCATAGATACCTCATCC
TGTCTACCTTTCAKGACCAABTACAGTTTTCATTGCAACGTCTGCCATCACAGTG
GGAATACCTATTTCCTCCGGAAGCAAGCAAACTTGAAGGAAATGTGCCTTAGTG
CTTTGGCCAACCTGACATGGCAGTCCCGAACACAGGATGAACATCCGAAGACAA
TGTTCTCCAAAGATAAGGATATTATACCATTTATTGATAAATACTGGGAGTGCAT
GACAACCAGACAGAGACCTGGGAAAATGACTTGGCCAAATAACATTGTTAAAA
CAATGAGTAAAGAAAGAGATGTATTCTTGGTAAAGGAACACCCAGATCCAGGC
AGTAAAGATCCAGAAGAAGATTACCCCAAATTTGGACTTTTGGATCAGGACCTT
AGTAACATTGGTCCTGCTTATGACAACCAAAAACAGAGCAGTGCTGTGTCTACT
AGTGGGAATTTAAATGGGGGAATTGCAGCAGGAAGCAGCGGAAAAGGACGAGG
AGCCAAGCGCAAACAGCAGGATGGAGGGACCACAGGGACCACCAAGAAGGCCC
GGAGTGACCCTTTGTTTTCTGCTCAGCGCCTTCCCCCTCATGGCTACCCATTGGA
ACACCCGTTTAACAAAGATGGCTATCGGTATATTCTAGCTGAGCCTGATCCGCA
CGCCCCTGACCCCGARAAGCTGGAACTTGACTGCTGGGCAGGAAAACCTATTCC
TGGAGACCTCTACAGAGCCTGCTTGTATGAACGGGTTTTGTTAGCCCTACATGAT
CGAGGCTTTGATAAAATTCAAGAGTTATTTGTATTTTGAGGAAAAAGACTTTGTG
GATAAAGCAGAGAAGAGCCTGAAGCAGACTCCCCATAGTGAGATAATATTTTAT
AAAAATGGTGTCAATCAAGGTGTGGCTTACAAAGATATTTTTGAGGGGGTTTAC
TTCCCAGCCATCTCACTGTACAAGAGCTGCACGGTTTCCCATTTACTTTGGACCA
TGSTTCAAGTATCCTCCGGAAGGATGYCACTTACCGCCCTATGANTGACATGGG

CTGGGGCGCCCTGGTAAAACACACCCTGGGCTGACGTCTTGTTATCACNTGGAN
ACACAAATTGG

SEQ ID NO:2

Predicted 305 amino acid protein product from the ORF at 413-1327 of SEQ ID NO:1

MDTQAGSVDEENGRQLGEVELQCGICTKWFTADTFGIDTSSCLPFXTXYSFHCNVC
HHSGNTYFLRKQANLKEMCLSALANLTWQSRTQDEHPKTMFSKDKDIIPFIDKYWE
CMTTRORPGKMTWPNNIVKTMSKERDVFLVKEHPDPGSKDPEEDYPKFGLLDQDL
SNIGPAYDNQKQSSAVSTSGNLNGGIAAGSSGKGRGAKRKQQDGGTTGTTKKARS
DPLFSAQRLPPHGYPLEHPFNKDGYRYILAEPDPHAPDPEKLELDCWAGKPIPGDLY
RACLYERVLLALHDRGFDKIQELFVF

SEQ ID NO:3

11 C-terminal amino acids specific to ASH2 encoded by SEQ ID NOs:1 and 5.

GFDKIQELFVF

SEQ ID NO:4

Exon found in previously described ASH2 genes, but not in the present genes

APQLKISDDRLTVVGEKGYSMVRASHGVRKGAWYFEITVDEMPPDTAARLGWSQP
LGNLQAPLGYDKFSYSWRSKKGTKFHQSIGKHYSSGYGQGDVLGFYINLPEDTETA
KSLPDTYKDK

SEQ ID NO:5

ASH2 cDNA--isoform SE 707

ATGGCGGCGGCAGGAGCAGGACCTGGCCAGGAAGCGGGTGCCGGGCCTGGCCC
AGGAGCGGTCGCAAATGCAACAGGGGCAGAAGAGGGGGAGATGAAGCCGGTG
GCAGCGGGAGCAGCCGCTCCTCCTGGAGAGGGGATCTCTGCTGCTCCGACAGTT
GAGCCCAGTTCCGGGGAGGCTGAAGGCGGGGAGGCAAACTTGGTCGATGTAAG
CGGTGGCTTGGAGACAGAATCATCTAATGGAAAAGATACACTAGAAGGTGCTG
GGGATACATCAGAGGTGATGGATACTCAGGCGGGCTCCGTGGATGAAGAGAAT
GGCCGACAGTTGGGTGAGGTAGAGCTGCAATGTGGGATTTGTACAAAATGGTTC
ACGGCTGACACATTTGGCATAGATACCTCATCCTGTCTACCTTTCATGACCAACT
ACAGTTTTCATTGCAACGTCTGCCATCACAGTGGGAATACCTATTTCCTCCGGAA
GCAAGCAAACTTGAAGGAAATGTGCCTTAGTGCTTTGGCCAACCTGACATGGCA
GTCCCGAACACAGGATGAACATCCGAAGACAATGTTCTCCAAAGATAAGGATAT
TATACCATTTATTGATAAATACTGGGAGTGCATGACAACCAGACAGAGACCTGG
GAAAATGACTTGGCCAAATAACATTGTTAAAACAATGAGTAAAGAAAGAGATG
TATTCTTGGTAAAGGAACACCCAGATCCAGGCAGTAAAGATCCAGAAGAAGATT
ACCCCAAATTTGGACTTTTGGATCAGGACCTTAGTAACATTGGTCCTGCTTATGA
CAACCAAAAACAGAGCAGTGCTGTGTCTACTAGTGGGAATTTAAATGGGGGAAT
TGCAGCAGGAAGCAGCGGAAAAGGACGAGGAGCCAAGCGCAAACAGCAGGAT
GGAGGGACCACAGGGACCACCAAGAAGGCCCGGAGTGACCCTTTGTTTTCTGCT
CAGCGCCTTCCCCCTCATGGCTACCCATTGGAACACCCGTTTAACAAAGATGGCT
ATCGGTATATTCTAGCTGAGCCTGATCCGCACGCCCCTGACCCCGAGAAGCTGG
AACTTGACTGCTGGGCAGGAAAACCTATTCCTGGAGACCTCTACAGAGCCTGCT
TGTATGAACGGGTTTTGTTAGCCCTACATGATCGAGGCTTTGATAAAATTCAAGA
GTTATTTGTATTTTGAGGAAAAAGACTTTGTGGATAAAGCAGAGAAGAGCCTGA
AGCAGACTCCCCATAGTGAGATAATATTTTATAAAAATGGTGTCAATCAAGGTG
TGGCTTACAAAGATATTTTTGAGGGGGTTTACTTCCCAGCCATCTCACTGTACAA
GAGCTGCACGGTTTCCATTAACTTTGGACCATGCTTCAAGTATCCTCCGAAGGAT
CTCACTTACCGCCCTATGAGTGACATGGGCTGGGGCGCCGTGGTAGAGCACACC
CTGGCTGACGTCTTGTATCACGTGGAGACAGAAGTGGATGGGAGGCGCAGTCCC
CCATGGGAACCCTGACCAGGTCCCTCTTTTCTGTCAAGGACTTTCTGGGAATAAT
ACTGGGGGTTTTGTTTTTGTTTTTGAACTGTCTCAAATGTTCTCCCAAAGATGCTA
AAAACACAGCCTCTCCTTTTAGCAAGTTAAAAGGCTGGGTAGGACTGCGGGAGA
CTGCCTGCCTTTCACCATTTTCTCCCCACTTCCAGTGACTGCTCTTATTTTGTGTA

CCATAAGCCAACAACCGCTGACTCCAGGATTGCATAAGCCCCCTGTGAAATCGG
TGCTGTACTGCATACCCTGCCAGCTGTGACTTGTTATCCTACTATATTTTCTAAG
GAGTGAATAATATTGTCCGAGTAACTAACTTATTTAAAAGACATTTCCTTCTGTG
GGCATTGACTGTATCCCACCTGTTTTCCAAGGAAATGGTAACCTGTTTCTGAGAA
CACCTGAAATCAATGGCTATACATTCCAAACCAATCTAAACGCTA

SEQ ID NO:6

amino acid sequence of isoform SE 707

MAAAGAGPGQEAGAGPGPGAVANATGAEEGEMKPVAAGAAAPPGEGISAAPTVEP
SSGEAEGGEANLVDVSGGLETESSNGKDTLEGAGDTSEVMDTQAGSVDEENGRQL
GEVELQCGICTKWFTADTFGIDTSSCLPFMTNYSFHCNVCHHSGNTYFLRKQANLK
EMCLSALANLTWQSRTQDEHPKTMFSKDKDIIPFIDKYWECMTTRQRPGKMTWPN
NIVKTMSKERDVFLVKEHPDPGSKDPEEDYPKFGLLDQDLSNIGPAYDNQKQSSAV
STSGNLNGGIAAGSSGKGRGAKRKQQDGGTTGTTKKARSDPLFSAQRLPPHGYPLE
HPFNKDGYRYILAEPDPHAPDPEKLELDCWAGKPIPGDLYRACLYERVLLALHDRG
FDKIQELFVF*GKRLCG*SREEPEADSP**DNIL*KWCQSRCGLQRYF*GGLLPSHLTV
QELHGFH*LWTMLQVSSEGSHLPPYE*HGLGRRGRAHPG*RLVSRGDRSGWEAQSP
MGTLTRSLFSVKDFLGIILGVLFLFLNCLKCSPKDAKNTASPFSKLKGWVGLRETAC
LSPFSPHFQ*LLLFCVP*ANNR*LQDCISPL*NRCCTAYPASCDLLSYYIF*GVNNIVRV
TNLFKRHFLLWALTVSHLFSKEMVTCF*EHLKSMAIHSKPI*TL

SEQ ID NO: 7

ASH2 cDNA for isoform SE 691

ATGGCGGCGGCAGGAGCAGGACCTGGCCAGGAAGCGGGTGCCGGGCCTGGCCC
AGGAGCGGTCGCAAATGCAACAGGGGCAGAAGAGGGGGAGATGAAGCCGGTG
GCAGCGGGAGCAGCCGCTCCTCCTGGAGAGGGGATCTCTGCTGCTCCGACAGTT
GAGCCCAGTTCCGGGGAGGCTGAAGGCGGAAAAGAAACCACTTTGCAGTGCCA
GACTGGAAGAAGTAACGGTCACTCTGAAAACAGGGTGGGAGAGCTGCCTCTCTT
TGAACCTCTCCCAGGACCAACTCTAACCCAGGGAGGCAAACTTGGTCGATGTAA

GCGGTGGCTTGGAGACAGAATCATCTAATGGAAAAGATACACTAGAAGGTGCT
GGGGATACATCAGAGGTGATGGATACTCAGGCGGGCTCCGTGGATGAAGAGAA
TGGCCGACAGTTGGGTGAGGTAGAGCTGCAATGTGGGATTTGTACAAAATGGTT
CACGGCTGACACATTTGGCATAGATACCTCATCCTGTCTACCTTTCATGACCAAC
TACAGTTTTCATTGCAACGTCTGCCATCACAGTGGGAATACCTATTTCCTCCGGA
AGCAAGCAAACTTGAAGGAAATGTGCCTTAGTGCTTTGGCCAACCTGACATGGC
AGTCCCGAACACAGGATGAACATCCGAAGACAATGTTCTCCAAAGATAAGGAT
ATTATACCATTTATTGATAAATACTGGGAGTGCATGACAACCAGACAGAGACCT
GGGAAAATGACTTGGCCAAATAACATTGTTAAAACAATGAGTAAAGAAAGAGA
TGTATTCTTGGTAAAGGAACACCCAGATCCAGGCAGTAAAGATCCAGAAGAAG
ATTACCCCAAATTTGGACTTTTGGATCAGGACCTTAGTAACATTGGTCCTGCTTA
TGACAACCAAAAACAGAGCAGTGCTGTGTCTACTAGTGGGAATTTAAATGGGGG
AATTGCAGCAGGAAGCAGCGGAAAAGGACGAGGAGCCAAGCGCAAACAGCAG
GATGGAGGGACCACAGGGACCACCAAGAAGGCCCGGAGTGACCCTTTGTTTTCT
GCTCAGCGCCTTCCCCCTCATGGCTACCCATTGGAACACCCGTTTAACAAAGATG
GCTATCGGTATATTCTAGCTGAGCCTGATCCGCACGCCCCTGACCCCGAGAAGC
TGGAACTTGACTGCTGGGCAGGAAAACCTATTCCTGGAGACCTCTACAGAGCCT
GCTTGTATGAACGGGTTTTGTTAGCCCTACATGATCGAGGCTTTGATAAAATTCA
AGAGTTATTTGTATTTTGAGGAAAAAGACTTTGTGGATAAAGCAGAGAAGAGCC
TGAAGCAGACTCCCCATAGTGAGATAATATTTTATAAAAATGGTGTCAATCAAG
GTGTGGCTTACAAAGATATTTTTGAGGGGGTTTACTTCCCAGCCATCTCACTGTA
CAAGAGCTGCACGGTTTCCATTAACTNTGGACCATGCTTCAAGTATCCTNCNGN
NGSATSYCACTTACCGCCCTATGAGTGACATGGGCTGGGGCGCCCTGGTAAAAC
ACACCCTGGGCTGACGTCTTGTTATCACNTGGANACACAAATTGGAATGGGGAA
GCCCCACTCCCCCCATGGGGGAAACCCTTGGAACCANGGTCCCTCCTTTTTCNTG
TTCCAAGGGA

SEQ ID NO: 8

amino acid sequence for isoform SE 691

MAAAGAGPGQEAGAGPGPGAVANATGAEEGEMKPVAAGAAAPPGEGISAAPTVEP
SSGEAEGGKETTLQCQTGRSNGHSENRVGELPLFEPLPGPTLTQGGKLGRCKRWLG

DRII*WKRYTRRCWGYIRGDGYSGGLRG*REWPTVG*GRAAMWDLYKMVHG*HI
WHRYLILSTFHDQLQFSLQRLPSQWEYLFPPEASKLEGNVP*CFGQPDMAVPNTG*T
SEDNVLQR*GYYTIY**ILGVHDNQTETWENDLAK*HC*NNE*RKRCILGKGTPRSR
Q*RSRRRLPQIWTFGSGP**HWSCL*QPKTEQCCVY*WEFKWGNCSRKQRKRTRSQ
AQTAGWRDHRDHQEGPE*PFVFCSAPSPSWLPIGTPV*QRWLSVYSS*A*SARP*PRE
AGT*LLGRKTYSWRPLQSLLV*TGFVSPT*SRL**NSRVICILRKKTLWIKQRRA*SRL
PIVR*YFIKMVSIKVWLTKIFLRGFTSQPSHCTRAARFPLTXDHASSILXXXXTYRPM
SDMGWGALVKHTLG*RLVITWXHKLEWGSPTPPMGETLGTXVPPFSCSKGTFLGG
N

SEQ ID NO:9

novel ASH2 amino acid subsequence from isoform SE 691

KETTLQCQTGRSNGHSENRVGELPLFEPLPGPTLTQGGKLGRCKRWLGDRII

WHAT IS CLAIMED IS:

1. A method of detecting cancer in a biological sample from an animal, the method comprising detecting the presence or absence of an ASH2 polypeptide or polynucleotide in said biological sample, wherein a diagnostic presence of said ASH2 polypeptide or polynucleotide indicates the presence of cancer in said biological sample.

2. The method of claim 1, wherein the detection is performed under conditions that would not detect said ASH2 polypeptide or polynucleotide in a sample that is not cancerous, and wherein a diagnostic presence comprises detecting any ASH2 polypeptide or polynucleotide in the biological sample.

3. The method of claim 1, wherein said diagnostic presence comprises at least a 2-fold increase in said ASH2 polypeptide or polynucleotide in the biological sample compared to a level expected in a sample from a control, cancer-free animal.

4. The method of claim 1, wherein said diagnostic presence comprises at least a 5-fold increase in said ASH2 polypeptide or polynucleotide in the biological sample compared to a level expected in a sample from a control, cancer-free animal.

5. The method of claim 1, wherein said diagnostic presence comprises at least a 10-fold increase in said ASH2 polypeptide or polynucleotide in the biological sample compared to a level expected in a sample from a control, cancer-free animal.

6. The method of claim 1, wherein the method comprises detecting a member selected from the group consisting of an ASH2 gene, ASH2 mRNA, and ASH2 polypeptides.

7. The method of claim 6, wherein the method comprises detecting an ASH2 polypeptide by immunoassay.

8. The method of claim 7, wherein the immunoassay comprises immunohistochemistry.

9. The method of claim 6, wherein the method comprises detecting the presence or absence of ASH2 gene amplification.

10. The method of claim 1, wherein said cancer is an epithelial cancer.

11. The method of claim 10, wherein said epithelial cancer is a cancer selected from the group consisting of breast cancer, lung cancer, colorectal cancer, prostate cancer, kidney cancer, stomach cancer, bladder cancer, ovarian cancer, and cancer of the gastrointestinal tract.

12. The method of claim 1, wherein said animal is a human.

13. The method of claim 1, wherein said animal is a mammal selected from the group consisting of non-human primates, canines, felines, murines, bovines, equines, porcines, and lagomorphs.

14. The method of claim 1, wherein said biological sample is selected from the group consisting of tissue biopsy, blood sample, buccal scrape, saliva, nipple discharge, and urine.

15. A method of detecting a cancer in a biological sample from an animal, the method comprising detecting the level of an ASH2 polypeptide or polynucleotide in said biological sample, wherein an increased level of said ASH2 polypeptide or polynucleotide in said biological sample compared to the level expected of a control sample from a normal, cancer-free animal indicates the presence of cancer in the biological sample.

16. The method of claim 15, wherein the difference between said increased level of said ASH2 polypeptide or polynucleotide and a level expected in a sample from a control, cancer-free animal is statistically significant.

17. The method of claim 15, wherein said increased level of ASH2 polypeptide or polynucleotide comprises at least a 2-fold increase in said ASH2 polypeptide or polynucleotide in the biological sample compared to a level expected in a sample from a control, cancer-free animal.

18. The method of claim 15, wherein said increased level of ASH2 polypeptide or polynucleotide comprises at least a 5-fold increase in said ASH2 polypeptide

or polynucleotide in the biological sample compared to a level expected in a sample from a control, cancer-free animal.

19. The method of claim 15, wherein said increased level of ASH2 polypeptide or polynucleotide comprises at least a 10-fold increase in said ASH2 polypeptide or polynucleotide in the biological sample compared to a level expected in a sample from a control, cancer-free animal.

20. The method of claim 15, wherein the method comprises detecting a member selected from the group consisting of an ASH2 gene, ASH2 mRNA, and ASH2 polypeptides.

21. The method of claim 15, wherein said cancer is an epithelial cancer.

22. The method of claim 15, wherein said animal is a human.

23. A method of monitoring the efficacy of a cancer treatment, the method comprising detecting the level of an ASH2 polypeptide or polynucleotide in a biological sample from an animal undergoing treatment for cancer, wherein a reduced level of said ASH2 polypeptide or polynucleotide in said biological sample compared to a level in a sample taken from the animal prior to, or earlier in, the treatment is indicative of efficacious treatment.

24. The method of claim 23, wherein said treatment is selected from the group consisting of chemotherapy, radiation therapy, surgery, hormone therapy, gene therapy, bone marrow transplantation, blood stem cell transplantation, and immunotherapy

25. The method of claim 23, wherein said cancer is an epithelial cancer.

26. The method of claim 23, wherein the method comprises detecting a member selected from the group consisting of an ASH2 gene, ASH2 mRNA, and ASH2 polypeptides.

27. The method of claim 23, wherein said animal is a human.

28. A method of treating cancer in an animal, said method comprising:

(i) detecting an ASH2 polypeptide or polynucleotide in a biological sample from said animal, wherein an increased level of said ASH2 polypeptide or polynucleotide in said biological sample from the animal indicates the presence of cancer in the biological sample; and

(ii) performing a cancer therapy in those animals having an increased level of said ASH2 polypeptide or polynucleotide.

29. The method of claim 28, wherein the method further comprises repeating step (i) after, or during, the performance of step (ii), wherein the efficacy of said cancer therapy is determined.

30. The method of claim 28, wherein the method comprises detecting a member selected from the group consisting of an ASH2 gene, ASH2 mRNA, and ASH2 polypeptides.

31. The method of claim 28, wherein said cancer is an epithelial cancer.

32. The method of claim 28, wherein said cancer therapy is selected from the group consisting of chemotherapy, radiation therapy, surgery, hormone therapy, gene therapy, bone marrow transplantation, blood stem cell transplantation, and immunotherapy.

33. The method of claim 28, wherein said animal is a human.

34. A method of decreasing the proliferation of a cell with an elevated level of ASH2, said method comprising contacting said cell with an inhibitor of ASH2.

35. The method of claim 34, wherein said cell is a cancer cell.

36. The method of claim 34, wherein said cell is a tumor cell.

37. The method of claim 35, wherein said cancer cell is an epithelial cancer cell.

38. The method of claim 37, wherein said epithelial cancer cell is a cancer cell selected from the group consisting of breast cancer cell, lung cancer cell, colorectal

cancer cell, prostate cancer cell, kidney cancer cell, stomach cancer cell, bladder cancer cell, ovarian cancer cell, and cancer of the gastrointestinal tract cell.

39. The method of claim 34, wherein said cell is a hyperproliferative cell.

40. The method of claim 36, wherein said cell is a metastatic cell.

41. The method of claim 34, wherein said inhibitor is selected from the group consisting of antisense polynucleotides, ribozymes, antibodies, dominant negative ASH2 inhibitors, and small molecule inhibitors of ASH2 activity.

42. An isolated nucleic acid comprising an ASH2 polynucleotide of at least 50 nucleotides in length, wherein:

(i) said ASH2 polynucleotide encodes an ASH2 polypeptide comprising an amino acid sequence that is at least about 70% identical to a sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:6, SEQ ID NO:8, and subsequences thereof; and

(ii) said ASH2 polynucleotide does not include a subsequence that is at least about 70% identical to SEQ ID NO:4.

43. The nucleic acid of claim 42, wherein said ASH2 polynucleotide encodes a full-length ASH2 polypeptide.

44. The nucleic acid of claim 42, wherein said ASH2 polynucleotide comprises a nucleotide sequence that is at least about 70% identical to a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:5, SEQ ID NO: 7, and subsequences thereof.

45. The nucleic acid of claim 42, wherein said ASH2 polynucleotide comprises a nucleotide sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:5, and SEQ ID NO:7.

46. An isolated ASH2 polypeptide of at least 20 amino acids in length, wherein:

(i) said ASH2 polypeptide comprises an amino acid sequence that is at least about 70% identical to a sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:6, SEQ ID NO:8, and subsequences thereof; and

(ii) said ASH2 polypeptide does not include a subsequence that is at least about 70% identical to SEQ ID NO:4.

47. The ASH2 polypeptide of claim 46, wherein said polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:6, and SEQ ID NO:8.

48. The ASH2 polypeptide of claim 46, wherein said polypeptide is a full-length ASH2 polypeptide.

49. An isolated nucleic acid that comprises an ASH2 polynucleotide of at least 50 nucleotides, wherein:

(i) said ASH2 polynucleotide encodes an ASH2 polypeptide comprising an amino acid sequence that is at least about 70% identical to a sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:6, SEQ ID NO:8, and subsequences thereof; and

(ii) said ASH2 polypeptide comprises a subsequence that is at least about 70% identical to a sequence selected from the group consisting of SEQ ID NO:3 and SEQ ID NO:9.

50. The nucleic acid of claim 49, wherein said ASH2 polynucleotide encodes a full-length ASH2 polypeptide.

51. The nucleic acid of claim 49, wherein said ASH2 polynucleotide comprises a nucleotide sequence that is at least about 70% identical to a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:5, SEQ ID NO:7, and subsequences thereof.

52. The nucleic acid of claim 49, wherein the ASH2 polynucleotide has a sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:5, and SEQ ID NO:7.

53. An ASH2 polypeptide of at least 20 amino acids in length, wherein:

(i) said ASH2 polypeptide comprises an amino acid sequence that is at least about 70% identical to a sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:6, SEQ ID NO:8, and subsequences thereof; and

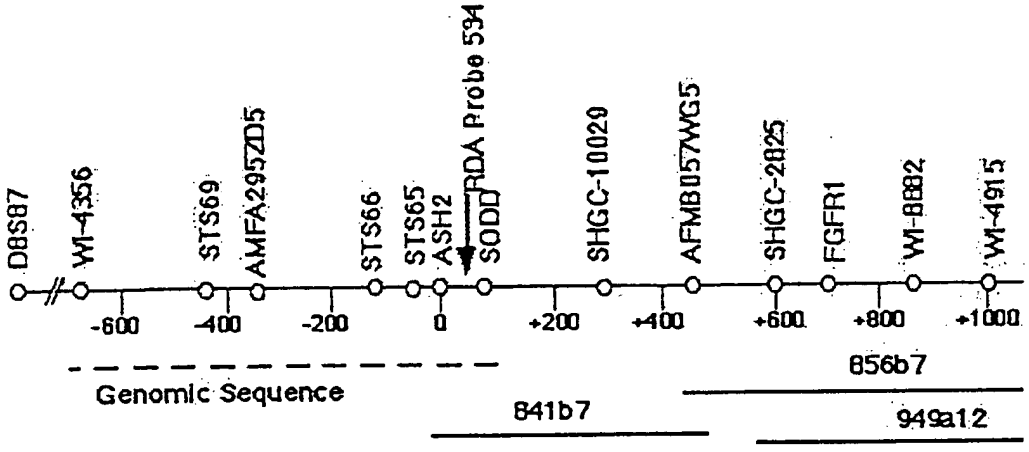
(ii) said ASH2 polypeptide comprises a subsequence that is at least about 70% identical to SEQ ID NO:3 or SEQ ID NO:9.

54. The ASH2 polypeptide of claim 53, wherein said polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:6, and SEQ ID NO:8.

55. The ASH2 polypeptide of claim 53, wherein said polypeptide is a full-length ASH2 polypeptide.

56. An isolated antibody that specifically binds to the ASH2 polypeptide of claim 46 or claim 53.

57. The antibody of claim 56, wherein said antibody binds specifically to a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:3 and SEQ ID NO:9.



Fold amplification:

87-425	2	1	4	4	9	7	12	5	6	2	3	6	6	2
88-523	4	4	2	3	9	9	7	5	9	5	4	10	5	4
88-688	2	10	3	7	10	9	11	8	6	7	8	8	10	8
89-249	1	4	1	2	4	3	5	3	3	5	3	3	3	1
96-201	2	2	1	2	3	5	5	4	2	1	3	3	1	2
96-76	1	5	1	3	5	7	7	1	3	3	5	8	7	3

Figure 1